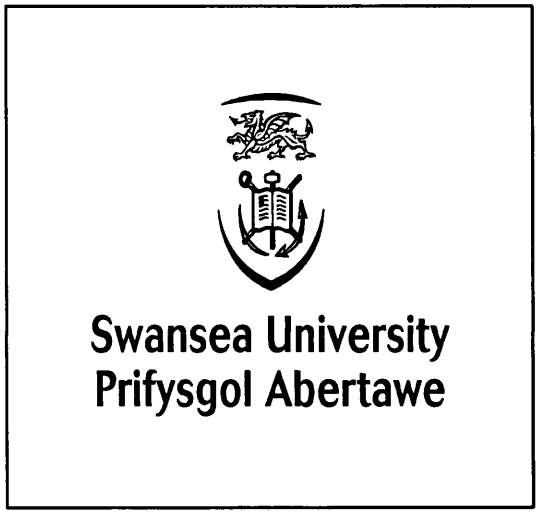
Optimising the refolding conditions of a recombinant protein.

Eldawi, Duha Ghaboush

OPTIMISING THE REFOLDING CONDITIONS OF A RECOMBINANT PROTEIN

BY

DUHA GHABOUSH ELDAWI



THESIS SUBMITTED TO THE UNIVERSITY OF WALES IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF PHILOSOPHY

SWANSEA UNIVERSITY

2008

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Summary

The objective of this project is to increase the yield and the quality of *in vitro* refolded proteins over-expressed in *E. coli* as insoluble inclusion bodies by optimising the refolding conditions. Therefore, protein refolding additives and regimes like (Glycerol, Gnd-HCl, Dilution, Temperature, and Triton X-100) were investigated to determine their effect on tackling the problems that initiate aggregation in refolding media.

The study identified Glycerol as an unsuitable aid in the refolding of Con A due to the binding effect it possessed towards the affinity dextran. Moreover, recommended the use of 0.5 M Gnd-HCl concentration as the initial point that guarantee suitable refolding yields. The rapid and continuous dilution protocols were proved to be more suitable than the other exploited refolding techniques in improving the refolding yields of Con A. However, rapid dilution was the preferable method to use since it is simple and reproducible. The 30X dilution factor was considered the optimum to enhance the refolding yields.

A gradual warm up step with 2 hours ice incubation period instead of 1 hour combined with the above mentioned technique ensured the recovery of most of the refolded Con A. Triton X-100 was unsuitable to aid the refolding of Con A due to its inhibitory effects on essential hydrophobic associations. However, its use as a tool was considered a breakthrough for estimating the time-course of refolding of the lectin.

Following the resultant optimised strategy the renaturation yield was improved in our laboratory from a mere 12 to 33 mg/l culture i.e. a 2.75-fold increase, providing that more detailed experiments are conducted to improve the use of Triton X-100 as a tool method using purified proteins with accurately known concentration values.

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ..... (candidate)

Date 30/06/2008

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ..... (candidate)

Date 30/06/2008

Statement 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate)

Date 30/06/2008.

Table of Contents

TO MOHAMED & MY MOTHER

Acknowledgement

First I would like to extend my thanks to my supervisor Dr. D. H. Jones for his kind support and patience and also for the genuine scientific knowledge he provided in the most accurate and professional manner. I really appreciate his advice and help.

My grateful thanks are extended to Mr. Jim Young who was always there for me ready to help. His friendship is something I will remember for the rest of my life.

Also I'm grateful to my friends and colleagues in the biological sciences laboratory who were very kind and helpful to me.

My special thanks are also extended to my family back in Sudan whom I know had missed my presence with them.

I owe a lot to my husband (Mohamed) who prepared the final lay out of the manuscript, without his encouragement and support I was really lost. My deepest love is for him.

List of Figures

List of Tables

Chapter One

General Introduction

1.1 Expression System

1.1.1 Introduction

The production of pharmaceutical proteins through the expression of cloned genes in prokaryotic cells is widely exploited both in industry and research (Clark, 1998). The production of recombinant proteins depends on: expression levels, cell growth characteristics, post-translational modifications, and biological activity of the protein under investigation. The metabolic functions of the cell have to be controlled in a very precise manner because of the need to coordinate their activities, and to prevent the wasteful synthesis of unnecessary materials, or the accumulation of toxic products. This control is achieved at two main levels: regulation of the level of production of specific enzymes, and control of the activity of those enzymes. The activity of an enzyme, which essentially means its affinity for a specific substrate and the rate at which it carries out the reaction, is genetically determined by its structure, as well as being influenced by the surrounding conditions. The susceptibility of an enzyme to feed back inhibition (or activation), and other properties such as temperature stability, are also determined by its structure.

Analysis of how the informational events accelerates from the structure of a gene to the activity of the enzyme as the final product; reported a series of factors that can mark the overall regulation of that activity (Dale and Park, 2004). In summary, these are:

1-The number of copies of the gene:

The level of product depends on the number of copies.

2-The efficiency with which the gene is transcribed:

Many factors affect the amount of *m*RNA and the initiation of transcription by RNA polymerase is playing a leading role in that.

a-The stability of *m*RNA:

The rate of production and the length of time each molecule exists in an active state in the cell are important effectors in determining the amount of specific *m*RNA. Most bacterial *m*RNA persists in a timid time, with a half-life of about 2 minutes. One of the important features in the bacterial adaptation to changes in their environments is due to the stability of bacterial *m*RNA, although, some bacterial *m*RNA species are more stable than others. The *m*RNA encoded by some bacteriophages is clearly more stable than is usual in bacterial cells. Other forms of RNA (*r*RNA, *t*RNA) are also to some extent

more stable, which can be related to the high degree of secondary structure owned by these molecules.

b- The efficiency with which the *m*RNA is translated into protein:

This will be influenced by the efficiency of inhibition, and also by factors that affect the rate at which the ribosome travel along the *m*RNA.

c- The stability of the protein product:

As with the *m*RNA, the amount of protein indicates both its rate of production and its stability. Different proteins varying their stability to a noticeable degree, as might be expected from their different functions: a protein that forms part of a cellular structure is likely to be more stable than one that transmits a signal for switching on a transient cellular event.

d- Post-translational effects:

This consists of a series of events such as protein folding which is necessary for conversion of polypeptide chains into biologically active forms of structure, as well as covalent modifications that can affect the activity of the protein. Phosphorylation is especially important as a mechanism for regulating the activity of specific proteins.

1.1.2 *Escherichia Coli*

In the present study mature Concanavalin A is being expressed in *Escerichia coli* because of its many advantages as a valuable host according to (Hannig and Makrides, 1998; Makrides, 1996; and Goeddel and Gold, 1990). Therefore, a brief description of the organism and its characteristic is vital to enlighten the role it played in this part of the project. *E. coli* a gram-negative, non-pathogenic, short, motile, rod-shaped organism always ‹ 1 micrometer in length present naturally as an intestinal fauna organism that is not normally available in water or soil. The organism has a high capability of multiplication which enabled it to be the favourite choice when considering laboratory research. Simple culture media containing ions like (Na^+, K^+, Mg^{2+}, Ca^{2+}, NH_4^+, Cl^-, HPO_4^{2-}, and SO_4^{2-}), trace elements and glucose are the best attributes for good *E. coli* growth production. The logarithmic phase of *E. coli* growth rate is about 22 minutes at an incubation temperature of 37°C. The microorganism is a non selective bacterium that can be grown either aerobically or anaerobically, although, the optimal production always occurred aerobically. The production of recombinant proteins using *E. coli* as a host require specific complex media that is rich in amino acids, vitamins, salts, trace

elements, and glucose. Moreover, special equipment like certain shaped flasks that can allow for maximum aeration and also shaking incubators at optimum temperatures are essential parts in the large scale production of recombinant proteins. *E. coli* is the most frequently used prokaryotic expression system for the high-level production of heterologus proteins. However, despite its many advantages, the efficient expression of different genes in *E. coli* wasn't a routine matter, as the structural features of different genes and their transcribed *m*RNAs preclude the adoption of a generally applicable expression method.

1.1.3 Essential Components of Expression Vectors

Hanning and Makrides (1998) described the prokaryotic expression vector as a set of homogeneous genetic factors that influence both transcriptional and translational aspects of protein production. Furthermore, the vector copy number is dictated by the origin of replication (*ori*) and the selected vector through the inclusion of antibiotic-resistance gene.

1.1.3.1 Gene Copy Number

Most genes on the bacterial chromosome are present as single copies (with a few notable exceptions such as the genes for *r*RNA and for many of the specific *t*RNA species); gene copy number is not therefore an important method of control for most of the normal metabolic activities of a bacterial cell. However, it becomes important, when consider plasmid-mediated characteristics, particularly with reference to the cloning and expression of heterologus DNA. Some plasmids are present within the cell in very high numbers (running to thousands of copies), and this is reflected in enhanced level of expression of the genes they carry.

1.1.3.2 Transcriptional Control

1.1.3.2.1 Promoter

The promoter is a site on DNA to which RNA polymerase can bind and begin transcription. It is composed of a hexa-nucleotide sequence that is positioned approximately 35bp upstream of the transcription initiation base separated by a short spacer from another hexa-nucleotide sequence. Different types of promoters can affect

the expression of genes in *E. coli;* therefore, selection of the suitable promoter depends according to Goldstein and Doi (1995) and Yansura and Henner (1990), on; The promoter strength where it should be able to produce more than10-30% of the total cellular protein as expressed protein, and represent the minimum level of basal transcription. Promoters should be able to produce a noticeable effect in a straightforward and inexpensive way where they are used as thermal or chemical inducers. Isopropyl-B-D-thiogalactopyranoside (IPTG) is an example of effective inducers of the powerful hybrid *lac* (De Boer et al.,1983) and *trc* (Brosius et al., 1985) promoters although their usage in industry is quite restricted because of high costs and toxicity.

.

1.1.3.3 Transcriptional Terminator

A promoter during transcription can terminate its function, unless a transcription terminator is placed up-stream of the promoter, moreover, this transcription terminator as indicated by (Nishihara et al., 1994) when present up-stream has a negative effect on background transcription. Continued transcription from strong promoters into the replication region can destabilize plasmids, owing to overproduction of the recombinant protein, which involved in the control of plasmid copy number. In addition, transcription terminators enhance *m*RNA stability and can substantially increase the level of protein production as reported by Hayashi and Hayashi (1985).

1.1.4 Protein Synthesis

The translation of the *m*RNA proceeds near its 5` terminal with the formation of the corresponding amino terminal of the protein molecule. The message is read from 5` to 3`, concluding with the formation of the carboxyl terminal of the protein. Again the concept of polarity is apparent. The transcription of a gene into the corresponding *m*RNA or its precursor first forms the 5` terminal of the *m*RNA molecule. In prokaryotes, this allows for the beginning of *m*RNA translation before the transcription of the gene is completed. In eukaryotic organisms, the process of transcription is a nuclear one; *m*RNA translation occurs in the cytoplasm. This precludes simultaneous transcription and translation in eukaryotic organisms and makes possible the processing necessary to generate mature *m*RNA from the primary transcript-*hn*RNA.

1.1.4.1 Translational initiation

Environment of ribosome binding site (RBS) sequence with adenine and thymidine residues, the mutation of specific residues and the use of translationally coupled systems, are procedures that have been improved to minimize the initiation for secondary formation at the 5`end of the transcript. In addition, the optimised characteristics at the 5`end of the *m*RNA transcript are known as the main factors in providing an efficient way of *m*RNA translational initiation, (Makride, 1996).

1.1.4.2 Translational enhancers

Translational enhancers are sequence components from *E. coli* and bacteriophages include: a sequence from the *T7*-phage gene-10 leader; U-rich regions in the 5` untranslated region (UTR) of certain *m*RNA, such as the *E. coli* at pET gene; and the `down stream box` located directly down stream of the start codon in *T7* genes. Little is known about their mode of action as these sequences do not function as universal translational enhancers in *E. coli* (Andrews et al., 1996).

1.1.4.3 Translational terminators

*m*RNA translation is terminated by stop codons and *E. coli* expression vectors usually contain all three stop codons that stop the ribosome 'skipping'. *E. coli* represents a sensitive system towards the UAA codon; however, the translational-termination efficiency is more developed in the context of the tetra- nucleotide UAAU (Poole et al., 1995).

1.1.5 THE pET System

The pET System is designed as a strong, potent scheme especially developed for the cloning and expression of recombinant proteins in *E. coli* (Figure 1.1). Selected genes are cloned in pET plasmids under the influence of strong bacteriophage T7 transcription and (optionally) translation signals; expression is initiated by offering a source of T7 RNA polymerase to the host cell which is very sensitive and active that, when fully

induced, almost all of the cell's resources are converted into the selected gene expression; more than 50% of the total cell protein qualify for the product of interest in a few hours after induction. Although this system is extremely powerful, it is also possible to decrease the concentration of effectors for a more refined expression level. Decreasing the expression level can lead to improvement of the soluble yield of some selected proteins. Another vital advantage of this system is its ability to obtain the selected genes transcriptionally silent in the un-induced state. Selected genes are initially cloned using hosts that do not contain the T7 RNA polymerase gene, therefore, avoid plasmid instability due to the production of proteins that are potentially toxic to the host cell. The expression of the selected genes present in a non-expression host, might be facilitated either by infecting the host with λCE6, a phage that carries the T7 RNA polymerase gene under the control of the λ*p* and p_1 promoters, or by transferring the plasmid into an expression host containing a chromosomal copy of the T7 RNA polymerase gene under *lac*UV5 control. The latter case, expression is initiated by the addition of IPTG to the bacterial culture. Although in some cases (e.g., with innocuous target proteins) it may be possible to clone directly into expression hosts, though this is not recommended as a general approach. Two types of T7 promoter and several hosts that differ in their potency of inhibiting basal expression levels are available, providing great flexibility and the ability to optimise the expression of different types of selected genes.

The pET vectors were originally described by (Studier and Moffatt, 1986; Rosenberg et al., 1987; Studier et al., 1990). The newest pET vectors developed at Novagen offer improved characteristics that allow for easier cloning, identification, and purification of the selected proteins.

Transcription vectors: are designed for expression of the selected genes that already carry their own prokaryotic ribosome binding site and AUG start codon, therefore, the transcription vector pET-26(+) is used in the present project.

Translation vectors: composed of the highly efficient ribosome binding site from the phage T7 major capsid protein and used for the expression of selected genes without their own ribosomal binding site. The selective markers amp (ampicillin resistance) and Kan (kanamycin resistance) are available with the pET vectors. Both types of selection are widely exploited, but specific guidelines should be followed when using vectors carrying the β -lactamase gene. While ampicillin resistance is commonly used for selection in different cloning vectors, kanamycin resistance may be favoured under

certain conditions, such as: protein expression in laboratories requiring GMP standards and sub-cloning target genes from other ampicillin-resistant vectors. Ampicillin selection diminishes in cultures because secreted β -lactamase and the low pH that accompanies bacterial fermentation both degrade the drug. Practical schemes to overcome the loss of drug resistance are to replace the medium with fresh ampicillin-containing media or to use the related drug, carbenicillin, which is less sensitive to low pH in the absence of a source of T7 RNA polymerase. Background expression is minimized with the lack of T7 RNA polymerase because the host RNA polymerases do not develop from T7 promoters and the cloning sites in pET plasmids are in regions that are weakly transcribed (if at all) by read-through activity of bacterial RNA polymerase. Preferred bacterial hosts for cloning include the *E. coli* K12 strains Nova Blue, JM109, and DH5. These strains are perfect hosts for the start of cloning of target DNA into pET vectors and for developing plasmids because they are *recA– endA–* and have powerful transformation competence and good plasmid recovery yields. Nova Blue has the additional advantage of having a selectable F factor that allows helper phage infection and therefore the production of single stranded plasmid DNA for mutagenesis purposes. As for protein production, a recombinant plasmid is induced in an *E. coli* strain containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage λDE3, a lambda derivative that has the immunity region of phage *21* and carries a DNA fragment containing the *lacI* gene, the *lacUV5* promoter, and the gene for T7 RNA polymerase (Studier and Moffatt 1986; Novy and Morris, 2001). This fragment is inserted into the *int* gene, preventing λDE3 from integrating into or cutting from the chromosome without a helper phage. Once a λDE3 lysogen is formed, the only promoter known to lead transcription of the T7 RNA polymerase gene is the *lacUV5* promoter, which is inducible by isopropyl-b -D-thiogalactopyranoside (IPTG). Addition of IPTG to a growing culture of the lysogen initiates the T7 RNA polymerase production, which in turn transcribes the target DNA in the plasmid.

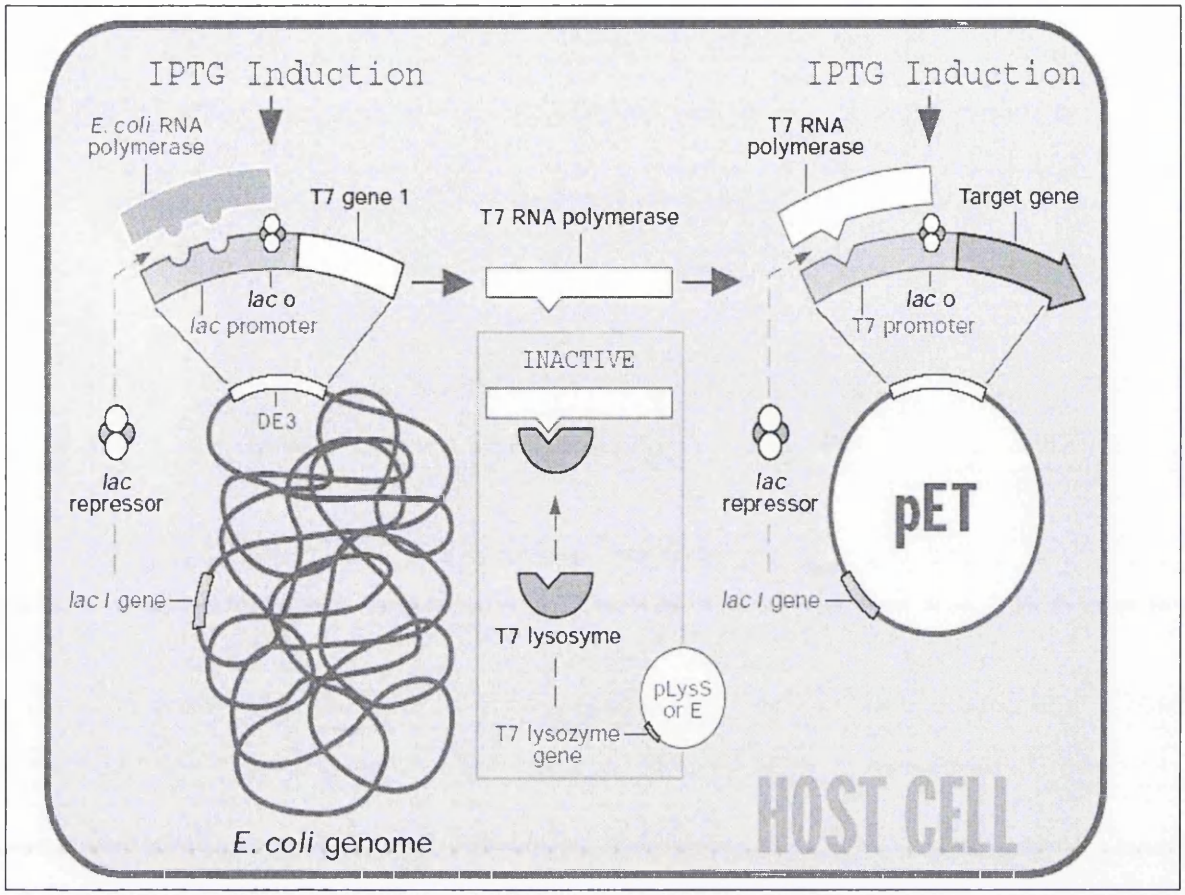


Figure1.1 Control elements of the pET system

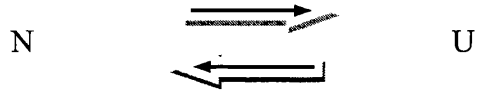
Showing the control elements of the pET system using *E. coli* as a chosen host where cloning starts in a non-expression host that does not contain T7 RNA polymerase gene
(Left side) and transferred into expression hosts containing a chromosomal copy of T7 RNA polymerase gene under *lac* UV5 control where expression induced by IPTG (right side)**(Reproduced from Novagen (2002-2003) Catalogue).**
Novagen pET system Tutorial. P84-91.

The pET 26(+) (Figure 2.1) used in the current project as the parent expression system for Concanavalin A and bacterial *ompA* signal was used in all constructs to direct the products to the periplasm of *E. coli*. The ompA signal derived from the pIN system was therefore continually used as it resulted in authentic N-termini of purified Con A before than the *pelB* originally provided in pET 26b (+) (Hanning and Makrides, 1998)).

1.2 Protein Folding

1.2.1 Introduction

Protein folding poses a puzzling question in science and technology and scientists in research and industry are facing many challenges to unmask the mysteries that surround it. Not all proteins can fold spontaneously or by random search through stable intermediates; therefore, it is clear that folding is a more complicated process that needs urgent attention. Although, scientists like Jaenicke and Rudolph, 1989; Creighton, 1994; Radford, 2000); have tried to predict the three-dimensional structure of a protein from its amino-acid sequence, more studies should be conducted until an answer can be found to that big question. Previously Levinthal explored the possibility of randomly calculating the number of the different conformations that arise from the denatured state of a protein and the results were largely disappointing since it appeared to be a time consuming task. This enormous difference between the calculated and the actual folding time is termed Levinthal's Paradox (Levinthal, 1968; Karplus, 1997).Proteins exist in marginal stability due to the difference in the free energy between the folded and unfolded states which can reach up to 10Kcal/mol where the average stabilization energy calculated for a 100-residue protein is only 0.1Kcal/mol and that is less than random thermal energy (RT=0.6Kcal/mol at room temperature). Most small proteins can be reversibly unfolded by raising the temperature or by adding a denaturing agent such as H^+, urea, or guanidine hydrochloride. Two conformational states are present in appreciable amounts a native state N (well-defined compact conformation) and unfolded or denatured state U (distorted or disordered) and they are in rapid equilibrium:



Thus, at the mid-point of the thermal or denatured states half the molecules are fully unfolded and the other half is fully recovered in native state. As known from Anfinsen, (1973); proteins are linear polypeptides that fold into active conformations depending on their primary structure, Levinthal, (1968); indicated that the folding process is not a simple but a complex issue since folding is a stepwise process. Other proteins exhibit an additional state which occurs under certain conditions that are represented by nearly the same secondary structure as the native protein described by Far Ultra-violet Circular Dichroism Spectra (CDS) hence: CDS distinguishes between ά helix, β sheet, and random coil structures. By contrast, the Near Ultra-violet CDS showed that many of the interactions of aromatic side chains were disrupted by acid treatments, and that the presence of a compact molecule is evident when the denaturing acids are used in low concentrations, this molecule is termed the molten globule. Moreover, other scientists have argued the presence of this intermediate illustrating that the denatured state cannot directly fold into the active form without collapsing into a form of an intermediate. The appearance of the molten globule is now a well documented area of research and our knowledge of it is dictated as compact polypeptides which are composed of the native secondary but not the tertiary structure. Molten globule formation is driven by the formation of helical stretches of the secondary structure. Many short peptides in water possess considerable helix- forming potential. The non-polar interior of molten globule stabilizes nascent helices by shielding their main- chain amides and carbonyl groups from competing with the water molecules, hence, states like dry, wet, and swollen molten globule have been observed relating the presence of water molecules and their interactions with the polypeptide chains (Pititsyn et al., 1995b; Peng and Kim, 1994; Redfield et al., 1994). Also many helices contain hydrophobic phases; the presence of these residues within the core of the molten globule also promotes helix formation. Thus hydrophobic collapse and acquisition of stable secondary structure are mutually reinforcing events in the formation of molten globules

Proteins can be denatured when subjected to unfavourable conditions that change their compact folded structure by distorting the native active conformation. Examples like chemical denaturants, thermal induction, acidic or alkaline media, and high pressure

stand as affecting factors in the folding process. As a result, when a protein is under suitable refolding conditions the molten globule appeared to be the first intermediate before the foundation of the fully folded native state, so Creighton (1997); postulated that the presence of MG is important for the protein to reach the fully folded native structure, but at the same time it has no role in enhancing or affecting the rapidity of the folding process.

The current work involves native proteins which are over-expressed in a bacterial host as inclusion bodies, subjected to denaturating agents, to reveal the optimum conditions for their refolding into native biologically active proteins. The following sections discuss in details: the inclusion bodies and the factors that might prevent the correct folding or mis- folding (aggregation).

1.2.2 Inclusion Bodies

1.2.2.1 Introduction

The production of inclusion bodies is the first part towards the accomplishment of a refolding process; therefore, over- expression of recombinant proteins extensively used recently is considered a major breakthrough in research and industry. Different system hosts either prokaryotes or eukaryotes are widely exploited, however, the most commonly explored ones are *E. coli* and yeast. The accumulation of recombinant proteins as cellular aggregates either in the cytoplasm or the periplasm of the host cell is known scientifically as the inclusion body formation. Inclusion bodies (IBs) are insoluble compounds that are, seen until recently as particles devoid of biological activity that upon purification IBs can give rise to active soluble proteins that are used extensively in many therapeutic areas. IBs can be solubilised through the use of chaotrops in high concentrations; denaturants like guanidine hydrochloride and urea are usually used with the aid of reducing agents such as α-mercaptoethanol to facilitate the solubilisation of IBs. Although, the use of these denaturants in high concentrations leads to random coil formation and exposure of the hydrophobic surfaces which are the main reasons of aggregation and, hence, loss of bioactivity (Singh and Panda (2005), the challenge is to gain bioactivity back from these inert particles, since IBs formation from cloned genes products is attributed to generation of an incorrectly transcribed or translated sequence or to secondary damage (Carrió and Villaverde (2002).

Solubilisation and purification of these IBs through suitable refolding techniques are the best sources for the optimum recovery of active proteins; moreover, the high levels of expression in *E. coli* can also be rated as one of the reasons that can add to the high levels of bioactive recovery. Certain advantages were achieved through IBs formation; (i) increased purity, (ii) effortless isolation (iii) less degradation of the expressed protein, (iv) added protection from proteolytic degradation, and (v) homogenised protein (with less contaminant). Furthermore, the work of Ventura and Villaverde (2006); about the protein quality of bacterial IBs indicated that they are transient reservoirs of aggregated polypeptides which can refold back into active protein under favourable cellular conditions. In the following parts of this section the various aspects of the IBs properties and their isolation, purification and solubilisation are identified.

1.2.2.2 Properties of Inclusion Bodies

Over-expression of recombinant proteins in a cellular host like *E. coli* or Yeast is the major reason for their accumulation as inclusion bodies. Inclusion bodies are described morphologically as refractile dense mis-folded or disordered polypeptide chains, located in the cytoplasm or the periplasm of the selected host. They are known as amorphous structures that can be seen using light microscopy (Georgiou et al. (1986); Bowden and Georgiou (1990); Baneyex (1999), and Georgiou and Valax (1999). Inclusion bodies differ in their sizes which can reach up to several μg; in addition they consist of more than 50% of the total cell protein in the form of recombinant protein deposits. Secondary structures of polypeptide chains and contaminants of outer membrane proteins are the principle aggregates that form those inclusion bodies (Reddy et al. (1998). Washing with detergents is one of the means to eliminate the membrane proteins from the inclusion bodies. Practical considerations to isolate the IBs from the surrounding contaminants in the form of mechanical methods as high pressure homogenisation, chemical, enzymatic treatments or a mixture of these methods together are used to disrupt inclusion bodies. The resultant disrupted IBs then subjected to low speed centrifugation for a specific period of time or filtrated to remove the soluble fraction of the protein from the insoluble one. Washing by EDTA, and low concentrations of denaturants and/or weak detergents such as Triton X-100 (see section 2.4.3), deoxycholate and octylglucoside (Georgiou and Valax (1999); De Bernardez

Clark et al. (1999); Batas et al. (1999); Muller and Rinas (1999); Huxtable et al. (1998); Fahey et al. (2000); Lee et al.(2000); Collins et al. (1999) is widely accepted to remove the stubborn membrane protein and available contaminants. Moreover, there are ongoing studies to reveal which of these methods are the most suitable for the isolation of the IBs scientists like Batas et al. (1999); reported that the size of the membrane pore is considered as an important aspect in achieving greater protein purity as this has led to the observation indicating that centrifugation takes advantage of the density differences between cell debris and inclusion bodies. Solubilisation is another way in the purification of IBs maintained by different methods generally used in research or industry. Many preferred methods of chemical and physical treatments like the use of denaturants, detergents or pH and pressure are accurately selected to guarantee efficient refolding yield. In the next section a brief discussion of some of these treatments will follow:

(1) Denaturant

Denaturing is considered the best way to solubilise the IBs, although this may completely or partially disrupt the protein structure, the method is extensively used. Solubilisation may be accomplished by the complete disruption of the protein structure (unfolding) or by the disruption of intermolecular interactions with partial unfolding of the protein. Minimum amounts of the denaturant are required to unfold a protein so the knowledge about the protein structure and the suitability of the denaturant is vital. The most commonly used denaturants are guanidine hydrochloride and urea. Plenty of the documented literature reported the use of relatively high denaturant (6–8 M) and protein (1–10 mg/ml) concentrations in the solubilisation step (De Bernardez Clark 2001). Guanidine hydrochloride is the current solubilizing agent used extensively in our laboratory (see section 2.4.4) due to its stability when compared to urea that decomposes to cyanate and ammonia as identified by Shirley (1995), and Pace et al. (1989). Special precautions must be taken when urea is used for inclusion body solubilisation, like addition of scavengers containing free amino groups to the solubilisation mixture to remove any cyanate which would otherwise react with the protein. Moreover, other denaturants may be used like the cationic surfactant cetyltrimethylammonium chloride (CTAC) Rudolf and Lilie (1996).

(2) Detergents

Detergents are widely used in the washing step to remove impurities so as to prevent further proteolytic degradation of the IBs protein (Georgiou and Valax 1999).

Moreover, they are also considered as solubilizing agents for the IBs. Extensively explored detergents are sodium dodecyl sulfate (SDS) and n-cetyl trimethylammonium bromide (CTAB) (De Bernardez Clark 2001). Since the use of detergents is not recommended in solubilisation as they may interact with the chromatographic procedure, the removal of these agents is important to achieve a suitable refolding yield. Cyclodextrins, linear dextrins or cycloamylose are compounds that can interfere with the detergent chemical structure by forming a mixture that can be removed from the refolding environment. During this project TritonX-100 is used in the washing step to purify ConA inclusion pellets (see section 2.4.6).

(3) (Reducing agents)

Such as dithiothreitol (DTT), dithioerythritol (DTE), and 2-mercaptoethanol (Lilie et al. (1998) and De Bernardez Clark (1998) are used in sufficient amounts to obtain maximum reduction of all cysteine residues present. These reducing agents are important factors in preventing non-native intra- and inter-disulfide bond formation in highly concentrated protein solutions at alkaline pH. It is important to mention that the model protein (Con A) used in the current project lacks disulfide bridges therefore the use of these reducing agents is out of question.

(4) pH and Temperature

pH has always been a key factor in the attainment of complete solubilisation, since at a certain pH the protein can unfold completely. However, at the same time extreme pH at both ends (very low and very high) could have an unfavourable effect on the protein in question because of irreversible chemical modifications. Scientists like Khan et al. (1998); Patra et al.(2000); and Hartman et al. (1999); investigated the use of high pH ($\geq$12) to solubilize growth hormones and pro-insulin. Some studies as reports from Patra et al.(2000) have explored different solubilisation methods for the recovery of human growth hormone from *E. coli* inclusion bodies and obtained the same results of solubilisation efficiencies when using 8 M urea, 6 M GdmCl, 1% SDS or 1% CTAB (all at pH 8.5) or 2 M urea (at pH 12.5).

(5) In situ solubilisation

Addition of the solubilizing agents directly to the fermentation medium at the end of the process is known as In situ solubilisation. The process has its advantages in the elimination of time-consuming and energy-consuming mechanical disruption methods

and also centrifugation or filtration. Although, the release of both proteinaceous and non-proteinaceous contaminants that may have to be removed before renaturation is attempted are found to be the main disadvantages (Georgiou and Valax (1999); and Futami et al. (2000).

1.3 Aggregation

Discussion of folding is not complete unless the obstacles encounter the process of correct folding are mentioned, therefore, this part will try to identify the different areas of the aggregation process and how scientists and researchers are trying to deal with the problems that arise from it. Aggregation is a process which occurs when intra- and intermolecular interactions driving protein folding are not correctly balanced during the renaturation or refolding process. Hydrophobic interactions are the major source of aggregation when the side-chains of a folding protein, or two interacting molecules come together there is a favourable increase in the entropy of the system, since, the solvent (water) molecules behaves in a random way breaching the ordered environment around the exposed hydrophobic surfaces. Moreover, unfavourable apolar-polar interactions are replaced with more favourable homo -type interactions leading to energy contribution. Because aggregation depends on protein concentration the way to tackle this problem is to reduce it by utilizing methods like dilution(see also section 1.4.2) which provides low concentrations, however, dilution has disadvantages which are represented in large volumes of buffers, large vessels and this expenditures lead to expensive budgets. Knowing the exact pathway of aggregation may be a good way to overcome this barrier since in industry to refold at low costs utilising easy methods, the pulse or step-dilution (Rudolph and Fischer, 1990) may be the perfect solution to refold at high concentrations. Another source of aggregation is the refolding intermediates; these intermediates are formed during the refolding pathway since they are mainly composed of secondary structure and lacking tertiary forms that interact together through their exposed hydrophobic surfaces making aggregation possible, hence, divert from the folding pathway. The "temperature leap tactic" method (see section 1.4.6) developed by Wetlaufer and Xie tried to offer a solution by sudden rising of temperature after cold incubation. The method utilize denaturant concentrations that are less than1M and indicate the formation of the first intermediate instantly from the beginning of the refolding process then its isomerisation into the second intermediate and so on until the

appearance of the fully folded native structure providing that suitable refolding conditions are maintained. Goldberg et al., (1996) revealed that aggregation is a non-specific phenomenon since incorrect disulphide bonding may not be the major cause of aggregation when they explored the nature of the different interactions responsible for aggregation during folding. On the contrary, (Speed et al., 1996) argued that larger aggregates could grow by a different mechanism involving non-specific interactions and their hypothesis was confirmed by (Maachupalli-Reddy et al., 1997) proved that the non-specific nature of the aggregation reaction as foreign proteins are likely to aggregate when folded in isolation. Hydrophobic interactions are the main cause of aggregation which is a fast process that occurs on the first minute of the refolding process while the aggregate size and not the concentration increase by time. Specific precautions are taken to reduce aggregation as (Katzav-Gozansky et al., 1996) identified mutations that disrupt the hydrophobic patches and the use of antibodies that bind them respectively; therefore, both strategies have the same goal in tackling the aggregation problem by preventing available hydrophobic interactions.

1.4 Protein Refolding

1.4.1 Introduction

The previous sections discussed some of the different aspects of folding and the problems that are facing the production of the correct folded proteins. Refolding is the bottleneck in the protein science since scientists are trying to find suitable methods and schemes to improve the renaturation yield, knowing that every protein is different and specific in its nature. Thus, refolding is a procedure that is in rapid improvement to facilitate the correct folding of recombinant proteins produced by the utilization of inclusion bodies, where, aggregation is the major problem that is challenging it and trying to overcome this barrier is a trial and error pathway. Different techniques are tried focusing on the efficiency of the refolding and the maximum renaturation yield. Moreover, the economical use of these techniques and their effortless adoption by either research or industry directed a discussion in the forthcoming section to explore the different procedures that scientists are trying to improve, and investigate their benefits and limitations.

1.4.2 Dilution (Rapid or Continuous)

In section (1.3) dilution was discussed as a method that to some extent improves the refolding yield that proceeds via first order kinetics while reduces the aggregation process following high order kinetics that is concentration dependant. In the current project dilution (Rapid or Continuous) is used as a basic method of refolding throughout the practical investigations as Middelberg (2002), and De Bernardez Clark (2001) reported an improvement in renaturation yield when refolded at low protein and denaturant concentrations and also for its simplicity, and economy as the solubilised protein is directly or continuously diluted into the refolding buffer. Although it is a perfect method when applied in small scales it has limitations at larger scale: the need for large volumes of refolding buffers and containers make it less economical and time-consuming in industry (Vallejo and Rinas, 2004). While this method increased the final renaturation yield and noticeably reduced aggregation Tsumoto et al (2003) added an insight into the rapid and continuous dilution when confirmed the role of the denaturant concentration in both methods. The Tsumoto groups' data indicated that the gradual increase in the denaturant concentration from the zero point to the final diluted concentration allows for a comparison between the start and final phases of refolding which is found to be insufficient to refold oligomeric proteins and, therefore, the application of rapid dilution is more appropriate in this instance. The previously mentioned reviews led the scientists to think of a method that can combine the simplicity and low cost of the dilution method with improved efficiency in obtaining excess renaturation yield as will be discussed in the next section as the pulse renaturation method.

1.4.3 Pulse Renaturation

A breakthrough improvement to the previous dilution method developed by (Rudolph and Fischer, 1990); depends on intervals or semi-continuous additions of soluble denaturant increasing the protein concentration in the media with each addition (Katoh and Katoh, 2000; Vallejo and Rinas 2004; Buchner et al. 1992; and Tershima et al. 1996). Furthermore, the method achieved high yields while refolding at high protein concentrations, thus, the knowledge of the final refolding kinetics, the volume of each addition corresponding to the maximum refolding yield (critical concentration above

which aggregation appeared), the maximum concentration of denaturant that gives rise to high yield, and the time interval between each addition which allow for sufficient refolding is crucial (De Bernardez Clark et al. 1999; Ming Li et al. 2003; and Singh and Panda, 2005).The method has advantage; the increase in the final renaturation yield, prevention of aggregation when the accumulation of the refolded particles is avoided, and also minimizing of the volumes of refolding buffers. On the other hand, the main drawbacks are the search for the suitable refolding buffers and the use of additive enhancers (Middelberg, 2002). The current studies have exploited the use of the pulse renaturation technique and step-wise additions and report the adaptation of the method to the model protein used.

1.4.4 Mixing

An important addition to both the methods of dilution and pulse renaturation used as part of the refolding techniques in the current studies, where, it keeps the concentration of both the denaturant and the protein constant as it helps in the formation of refolding intermediates through the rapid collapse of the protein (Tsumoto et al.2003; Jungbauer and Kaar 2007) reviewed the effortless use of mixers in devices composed of tankers and stirrers needed to apply the method in small scales, however, they reported the disadvantages of this methods due to inefficient mixing that can raise the possibility of aggregation and also the increase in time when applied in large scale production. A device relating mixing intensity to oscillatory Reynolds number (Re_O) is called the oscillatory flow reactor was developed by Lee et al., (2002, and 2001); overcome this barrier by a better mixing uniformity. Furthermore, Hohenblum et al. (2004) and Schlegl et al. (2005 a, and b); reported the use of continuous stirred tankers to eliminate denaturing reagents from the refolding medium to insure optimum refolding environment and also to reuse the aggregated protein to increase the overall yield.

1.4.5 Dialysis

Previously two types of dilution have been discussed and now in this section an exploration of the different opinions concerning the dialysis technique which is a modified method of the ordinary dilution will be held. As was mentioned before these techniques depend on the speed at which the denaturant concentration is decreased and also the length of time that the protein subunits spend during an intermediate denaturant concentration to reach the final fully refolded native structure. In contrast to the above methods, this technique differs in that; the change from denaturing to native buffer conditions occurs gradually. Thus, from what have been discussed in different parts of this introduction this will initiate the presence of folding intermediates which leads to aggregation (Tsumoto et al. 2003).As this procedure can be excellent in the refolding of some proteins it can be modified into a step-wise manner adopting a process that induce the unfolded protein first with high denaturant concentration then adopting middle concentration in the second step and finally low concentration in the third step (Kumagai and Tsumoto 1998; and Tsumoto et al. 1998) as an equilibrium should be reached at each step before starting the next one. Advantages of this procedure are illustrated in that at intermediate denaturant concentration misfolding or aggregation might occur but at the same time denaturant intermediate concentration allows for the reverse formation of the folded structure. Furthermore, the process can affect the refolding of multi-domain proteins as each domain may refold on its own since their folding using the one step dialysis causes aggregation.

1.4.6 Temperature

Incubation temperature is considered one of the potential factors that affect the purified yield either positively or negatively (Vallejo and Rinas, 2004; Buchner et al., 1992; West et al., 1998; Yoshii et al., 2000). Although, several scientists had pointed out the effect of low temperature on the slowdown of the refolding process (Yoshii et al., 2000) they also appreciated its role on the prevention of hydrophobic aggregation. While, De Bernardez et al. (1999) insisted on an incubation temperature of 15°C as the recommended starting point to refold new proteins, some scientists like (Xie and Wetlaufer, 1996) refolded proteins at low temperatures, to minimize aggregation, and

then rapidly raised the temperature to promote fast folding after the intermediates responsible for aggregation have been depleted ensuring high yields production. During the current project a study has risen from the above mentioned information to investigate the different refolding incubation temperatures starting from as high as 37°C to as low as –5.5°C, hence, exploring their effect on the refolding yield and aggregate formation.

1.4.7 Chemical aids

1.4.7.1 Guanidine hydrochloride (Gnd-HCl)

One of the strongest chaotropic denaturant that has been generally used to disrupt the ordered state of protein native structure. Its primary mechanism is to affect the energy of stabilization, hence affecting the entropy which enables order in the native structure. In addition, it also affects the enthalpy energy derives the non-covalent interactions, hydrophobic interactions, hydrogen bonds, and ionic bonds. Gnd-HCl works on the native state to disrupt the ordered conformation of the hydrophobic interactions misshaping a compact globular protein molecule to an open chain with loose core where, it requires lower amounts of denaturant to succeed to the fully folded protein. Although proteins have been successfully refolded from the denatured state, it may prove to be difficult to fold proteins from a partially folded state as the determinant step is the suitable minimum amount required to solubilise the protein and to allow for full bioactivity recovery in the refolding step as (Singh and Panda, 2005) observed the role played by low denaturant concentrations (1-2) M on improving the refolding yield. Other chaotropes such as alkyl-urea or organic co-solvents like carbonic acid amides may be used to improve *in vitro* folding (Rudolf and Lillie, 1996).

1.4.7.2 L-arginine

A popular additive used extensively in research recently is L-arginine as it has the advantage for partially folded chains where it enhances the solubility of the refolding intermediates (Vallejo and Rinas, 2004). As it was found that L-arginine-HC1 increases the yield of renaturation of human tissue-type plasminogen activator from *E. coli*

inclusion bodies, other proteins were tried also, among them antibody Fab fragments, single-chain immunotoxins, or casein kinase II. Rudolf and Lillie (1996) indicate the stabilisation effect of L-arginine with proteins other than cytochrome C which showed a slight destabilizing effect of L-arginine.

1.4.7.3 Glycerol

Glycerol an alcohol aid in the improvement of the refolding environment is one of the viscogenic cosolevents widely used in research because it has little effect on pH and does not alter the protein native configuration. Glycerol is considered as a stabilising agent that when used in high concentrations mixes with the mature folding intermediates during the refolding phase, thus, reducing the aggregation process (Meng et al., 2001). The current project investigates the role played by glycerol in the improvement of the refolding of Concanavalin A in a step to optimise its refolding conditions.

1.4.7.4 Detergents

Ionic and non-ionic detergents have been extensively used in the solubilisation of inclusion bodies; however, (Tandon and Horowitz, 1996), had reported the use of lauryl maltoside in the refolding of the globular protein rhodanese to suppress aggregation and improve the refolding capacity of the refolding buffers. Other detergents were also put under trial (1) Non-ionic detergents like Tween20, TritonX-100 and octylglucoside, (2) Ionic detergents like SDS, cetyltrimethylammonium bromide, and (3) Zwitterionic detergents, all were found to refold perfectly the globular protein rhodanese with the lauryl maltoside having the maximum recovery activity. The detergent properties of TritonX-100 used in the present project are further explored in a separate study to reveal its effect on the refolding of Concanavilin A.

1.4.7.5 Chaperones

Chaperones are a naturally occurring group of proteins aiding *in vivo* protein folding and protecting cellular proteins from different types of environmental conditions by minimising protein aggregation (Vallejo and Rinas, 2004).

Natural chaperones have also been used for *in vitro* protein refolding but their high cost, high concentration, and the need for their removal after the refolding process, restricts their wide application. Chaperones are not suitable large-scale refolding processes (Vallejo and Rinas, 2004).

Some detergents have been viewed as artificial chaperones because they may resemble the mode in which a natural chaperone acts through the formation of a protein complex. Refolding depends on the removal of the added artificial chaperone. The release of the folding-competent protein is initiated by the addition of cyclodextrins. Cyclodextrin polymer beads are used in the removal of the cyclodextrin-detergent complex by centrifugation and are also used in expanded-bed columns in semi-continuous refolding processes (Vallejo and Rinas, 2004).

1.4.8 Chromatographic Refolding

Chromatography is one of the important methods in achieving correct refolding of inclusion bodies. Three methods comprise the major or the most used techniques in research and industry: affinity chromatography [using immobilized metal chelates (IMAC) or bio-specific binding], size-exclusion chromatography (SEC), and ion exchange chromatography (IEC). Changes have been made to these processes since they were first used in the 1990s to improve the overall renaturation yield and the recycling of aggregates.

1.4.8.1 Affinity chromatography

1.4.8.1.1 Immobilized Metal Affinity Chromatography (IMAC)

A popular and efficient method used in the purification and refolding of tagged proteins depends greatly on the formation of high affinity complexes with immobilised divalent

metal ions. Refolding can be achieved as has been described by (Vincent et al. 2004) by a simple buffer exchange in a stepwise manner where the gradual reduction of the denaturant through use of a gradient to initialise folding of a protein is common practice. A major disadvantage of the technique when the adsorbed protein concentration is high at the point of pouring the refolded sample onto the column for aggregates can form and the elution yields become low. Although, redox- systems are required for proper refolding they have to be adjusted very carefully in IMAC by performing of a buffer exchange since interactions of some metal ions, like Cu^{+2} which forms mis-paired disulphide bonds can occur (Stempfer et al., 1996). Furthermore, due to the increase in hydrophobic forces at high salt concentration initiating the aggregation of folding intermediates optimisation of binding, washing, and refolding conditions should be considered a fact that was mentioned by (Jungbaur et al., 2004) in a study about chromatographic beds used in folding and refolding of proteins.

1.4.8.1.2 Bio-specific affinity chromatography

This technique depends on the ability of a protein to bind to an immobilised form of its natural ligand or substrate. In the case of Con A, which binds to mannose and glucose derivatives, a convenient immobilised ligand already exists as the glucose-polymer dextran. This method was first used for Con A purification directly from crude jack bean meal extracts (Concanavalin A as a tool) and was subsequently used for recombinant forms (Min et al., 1992). Sephadex G-75 is used as the pre-formed affinity matrix and methyl-α-D-mannopyranoside is used as bio-specific eluant since it is the mono-saccharide of highest affinity for Con A.

1.4.8.2 Size-exclusion Chromatography (SEC)

Technique depends on the size of the porous structures of the gel to differentiate between the aggregate forming folding intermediates and refolded subunits by their hydrodynamic radius and also to recycle these aggregates through the use of the delayed running front of the denaturant. The method is conducted by injecting the denaturant mixture of the protein onto the previously equilibrated columns and the concurrent elution of it with refolding buffers results in high yields compared with the common

used dilution techniques. Precautions should be taken using this technique as injecting of denaturant proteins of high concentration or eluting at high rates can lead to aggregate formation and reduces overall renaturation yields (Vallejo and Rinas (2004).

1.4.8.3 Ion Exchange Chromatography (IEC)

A method based on the binding of the unfolded denatured protein to the matrix. Introduction of the pH gradient increased the efficiency of the technique by improving the elution yield as observed by (Gu et al., 2002). Low salt concentrations represent a way of improving the ion exchange technique as it reduces the possibilities of hydrophobic interactions forming intermediate aggregates, thus, improving the refolding yield (Schlegl et al., 2003). However, the technique as described provided more protein purification than other chromatographic techniques for crude samples the while the size exclusion proved to be more adequate in the utilization of the partially purified ones.

1.5 Concanavalin A (Con A)

1.5.1 Introduction

Concanavalin is one of the plant lectin proteins separated by Sumner (1919) into two crystallisable fractions, Concanavalin B and Concanavalin A according to their solubility in NaCl. Since that time Concanavalin A has been known as having agglutination activity that precipitate erythrocytes, starchy mucins and glyco- proteins (Sumner and Howell, 1936). Its carbohydrate binding activity was further revealed by (Goldstein and Poretz, 1986) in showing specificity in binding saccharides containing α-D-mannose or α-D-glucose residues and also the possibility of binding oligo-saccharide sequences lacking these units. Concanavalin A of the jack bean (*Canavalia ensiformis*) was selected as a model in the present work to optimise the refolding conditions of a recombinant protein. The term Con A will be used through out the remaining of this manuscript to identify the protein. Con A was chosen on specific criteria presenting it as an ideal model since it owns agglutination and carbohydrate binding activities, and the reaction with dextrans affords an easy and inexpensive method for its purification through affinity chromatography (see section1.4.2.1). Furthermore, many types of bacteria agglutinate well with Con A marking it as an

invaluable tool in many areas of research. Con A interact actively with cell surface receptor carbohydrates, and has an impact on cells transformed by oncogenic viruses, hence inhibits growth of malignant cells in animals, and exhibits mitogenic activity. Naturally Con A exists in tetramer form Figures (1.2, a and b); however, following change in pH and temperature patterns it can exhibit monomer and dimer structures.

1.5.2 Chemical structure

The chemical structure of the lectin Con A appeared to be consisting of ellipsoidal domes of dimensions 42 X 40 X 39 Å (Becker et al., 1976). Each subunit of it consists of 237 residues and contains two metal sites necessary for saccharide binding. One site binds transition metal ions, manganese, while the other favours binding calcium Figures (1.3, a and b) showing skeletal and Stereo views of Con A protomer, and (Figure 1.4) showing the complete amino acid sequence of Con A.

1.5.2.1 Polypeptide Chains

The folded polypeptide chain is mainly formed of two anti-parallel pleated β sheets with the presence of a twisted sheet of seven strands extending all through the centre of the molecule and a semi- curved sheet of six strands which forms the back surface of the monomer. Any disturbance to the β sheet structure leads to the collapse of the cleavage between residues 118 and 119 and, therefore, the unfolding of the pleated sheet. The NH_2- COOH-termini and the metal binding site occur on the right of the twisted β sheet (Becker et al., 1976).

1.5.2.2 Metal Ion Binding Sites

The metal mainly (Manganese and calcium) binding sites are present in an octahedral structure having four protein groups and two water molecules as ligands (Hardman et al., 1982; Weisgerber and Helliwell, 1993; Naismith et al., 1993; Emmerich et al., 1994).The binding site of Mn^{2+} allows for the binding of other metals for example Co^{2+}, Ni^{2+} and Cd^{2+}. The protein ligand for $Ca^{2+}$are the side chains of Asp 10, Asn 14, Asp 19, and the carbonyl oxygen of Tyr 12. In addition, Mn^{2+} has the ligands which are

the side chains of Glu8, Asp 10, Asp 19, and His 24 all are bound to two water molecules (Concanavalin A as a tool) Figures (1.5 a and b) showing metal and ligand binding sites for the four Con A subunits.

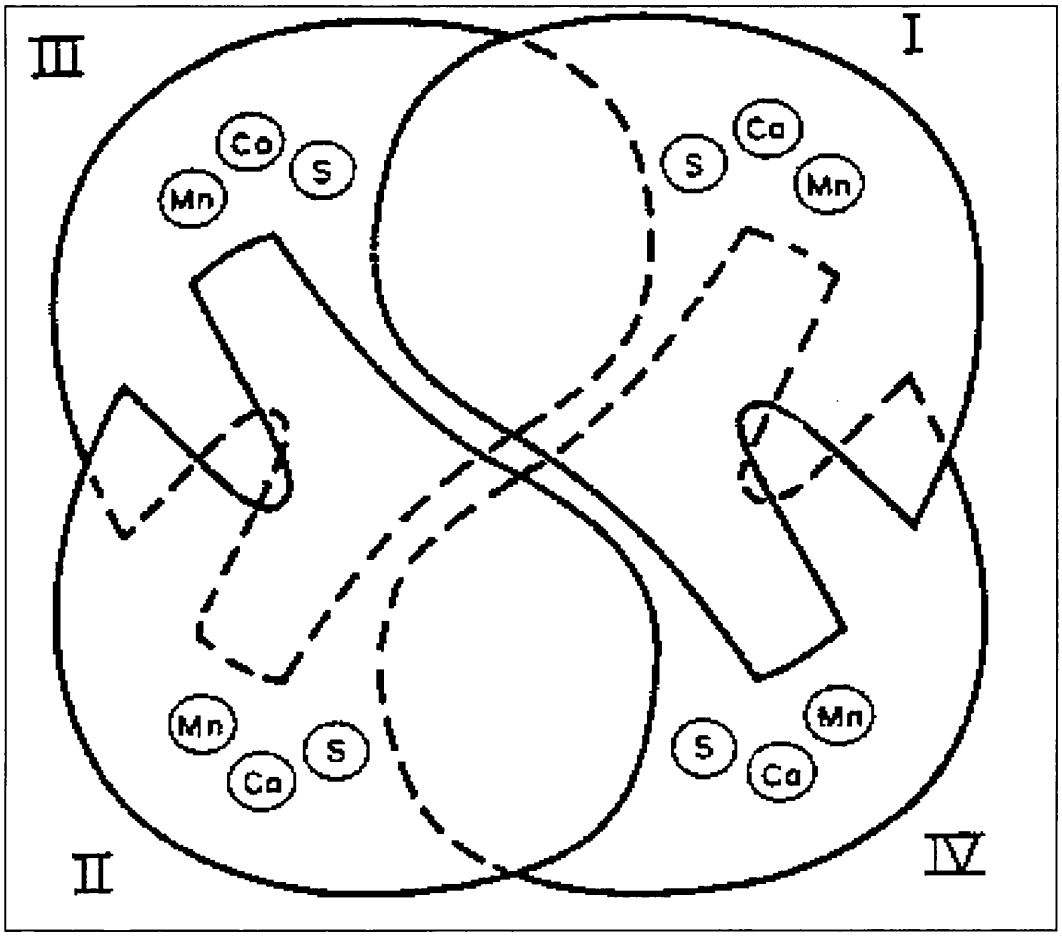


Figure1.2a Schematic diagram of the Con A tetramers

The four subunits showed in a heavy outline, are related by D_2 point symmetry. Binding sites for Ca^{2+}, Mn^{2+}, and specific saccharides are indicated by Ca, Mn, and S, respectively. Adapted from **(Reeke et al., 1975).** The covalent and three-dimensional structure of Con A. IV Atomic coordinates, hydrogen bonding, and quaternary structure. J. Biol. Chem. **250**: 1525-4

Figure 1.2b Ribbon representation of the Con A tetramer (crystal structure)

Polypeptide backbone for each subunit is drawn as a different coloured ribbon in yellow, blue, pink and green. The two divalent metal ions are shown as spheres: Mn^{2+} in yellow and Ca^{2+} in cyan. Bound monosaccharides (mannoside) are also shown as Van der Waals spheres with oxygen in red and carbon in black. Dimers existing in solution are formed by association between the yellow and blue subunits, and also by the pink and green subunits. Reproduced from (**Naismith et al., 1994**). Acta Cryst. **50**: 847 – 858)

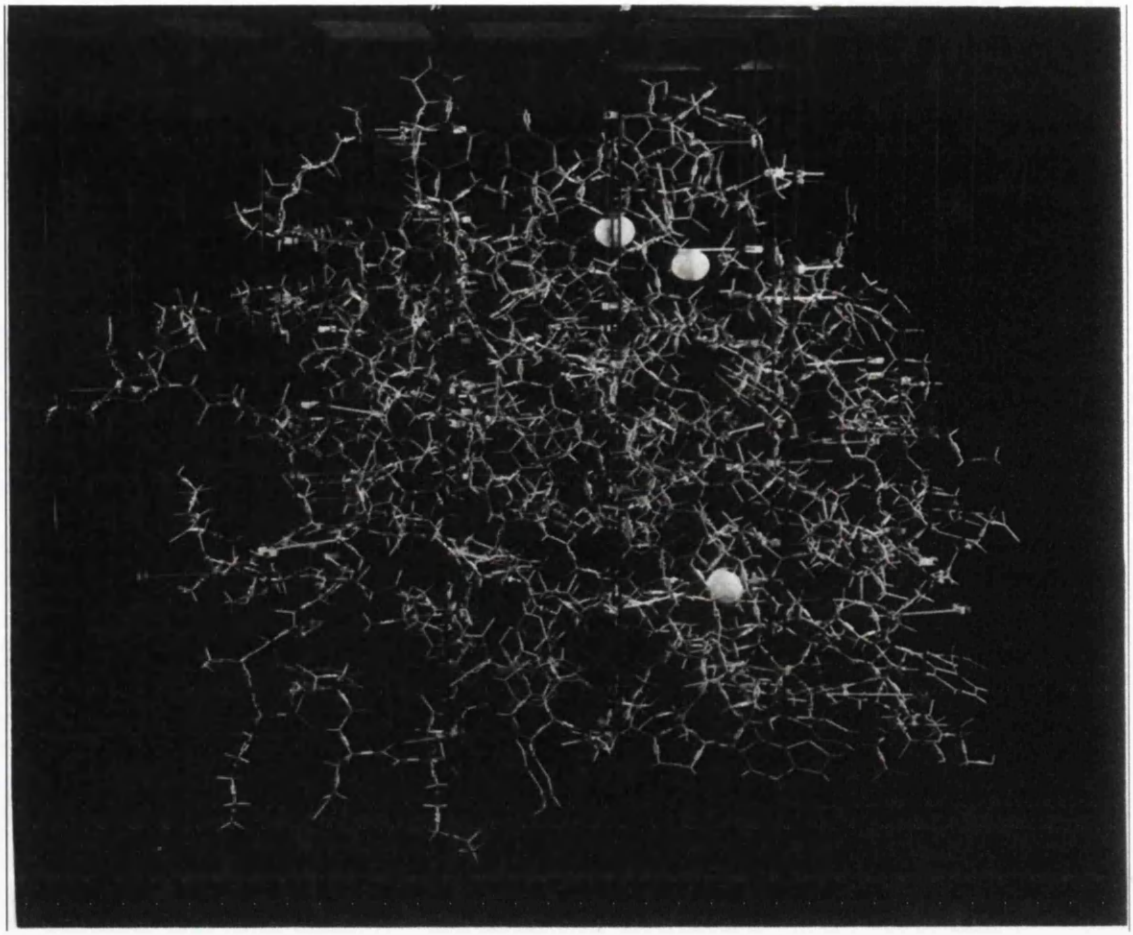


Figure 1.3a Skeletal model of the Con A protomer viewed down the z axis

[110] is horizontal, [110] is approximately vertical. The 2- fold axis relating members of an ellipsoidal dimer (second protomer not shown) is roughly perpendicular to the illustration at the bottom center. The metal atoms are represented by the two white spheres at the top of the model. The ß-IPG iodine-binding site is indicated by the sphere at the lower right. Sheet containing 14 hydrogen bonded chains, 7 from each monomer. Other interactions between halves of the ellipsoidal dimer apparently involve residues 87-88, 136-140, and 177-180. Reproduced from (**Edelman et al., 1972).** Proc. Nat. Acad. Sci. USA, **69**: No. 9, pp. 2580-2584.

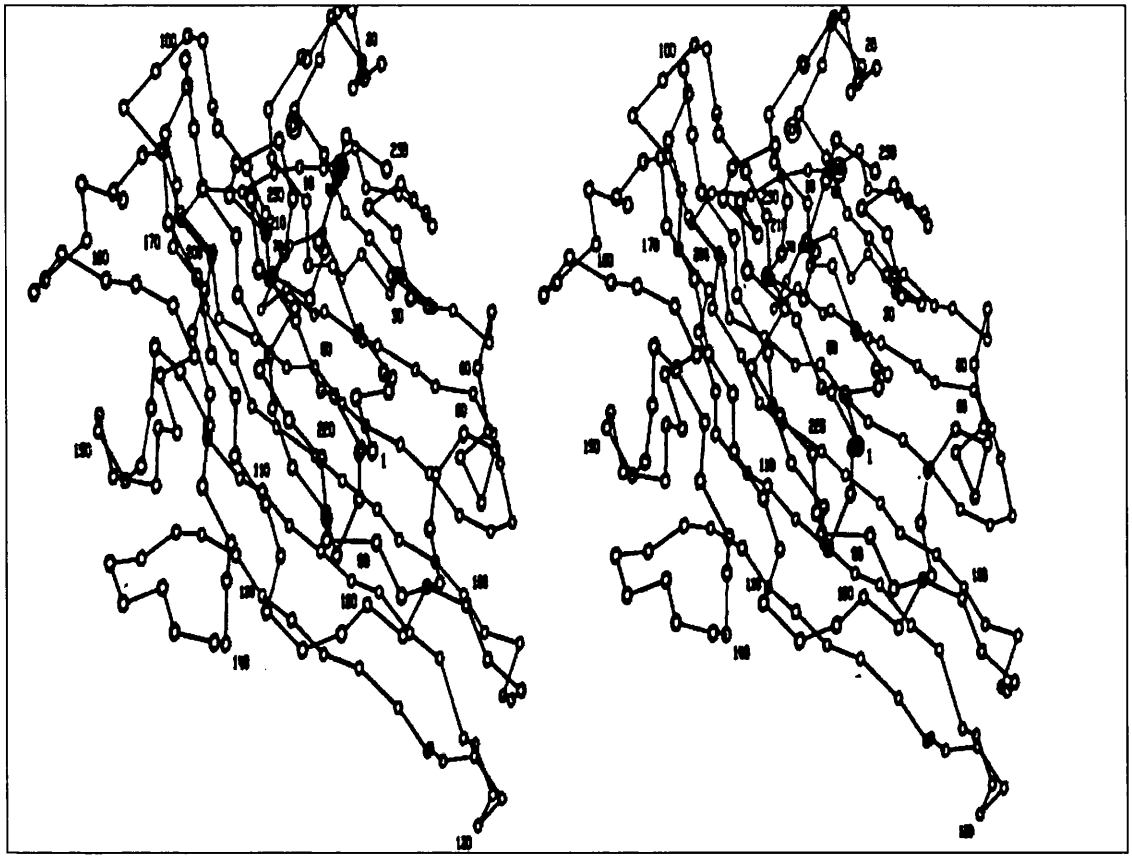


Figure 1.3b Stereo drawing of the a-carbon backbone of the Con A protomer

Figure 1.3b oriented as in Fig.1.3a the positions shown for residues 34 and 35 are tentative. Computer-drawn stereo figure was made by the program OR-TEP of Dr. Carroll Johnson. Reproduced from **(Edelman et al., 1972).** Proc. Nat. Acad. Sci. USA, **69**: No. 9, pp. 2580-2584.

1.5.2.3 Saccharide Binding Sites

The elaborate research of many scientists to investigate the ability of Con A to bind saccharides has given an insight into the true structure and function of its biological activities Figures (1.6 a and b) showing stereo views for the saccharide binding sites with networks of hydrogen bonds and metals. These studies identified specific mono-saccharides that can affect the biological activity of Con A, like D- glucose and D-mannose with the exclusion of other carbohydrates such as D- galactose. On the other hand, the presence of hydrophobic regions beside the saccharide binding sites strengthens some sugars as inhibitors than those without them (Goldstein et al., 1965; Poretz and Goldstein, 1970; and Goldstein et al., 1974). (Naismith et al., 1994; and Naismith and Field, 1996) described the Manoside binding site of Con A, also Kanellopoulos et al., (1996) observed the tri-Mannoside core found in N-linked glycans which explain the high affinity of Con A for these glycans and also reported a similar saccharide binding site for the 48-nitrophenyl-α-D-mannopyranoside complex Figures (1.7 and 1.8). Becker et al., (1976) indicated the interactions between the binding sites and the hydrophilic and hydrophobic regions in a study conducted on the inhibitor /3-oiodophenyl - n-glucopyranoside binding sub-sites.

In addition, the study concluded that the binding mode of the protein depends on whether it occurs in solution or in a crystalline state.

Farina and Wilkins (1980) also Troganis (1992) and other scientists have examined the role played by the aglycon group in binding and the determination of the various binding constants experimentally indicated their assistance in binding. Good quality crystals of Con A–saccharide complexes have been produced for a series of such compounds, which diffract at medium resolution, and X-ray diffraction data have been collected for several complexes. In a report by Kanellopoulos et al. (1996) the structure of the complex with 48- methylumbelliferyl-α-D-glucopyranoside (α-MUG) at 2.78 Å resolution has been described. However, Kanellopoulos et al. (1997) also revealed the crystal structures of complexes of Con A with 48-nitrophenyl-α-D-mannopyranoside (α-PNM) and 48-nitrophenyl-α-D-glucopyranoside (α-PNG), which indicates the higher affinity of Con A for Mannosides than for glucosides. However, it was recently

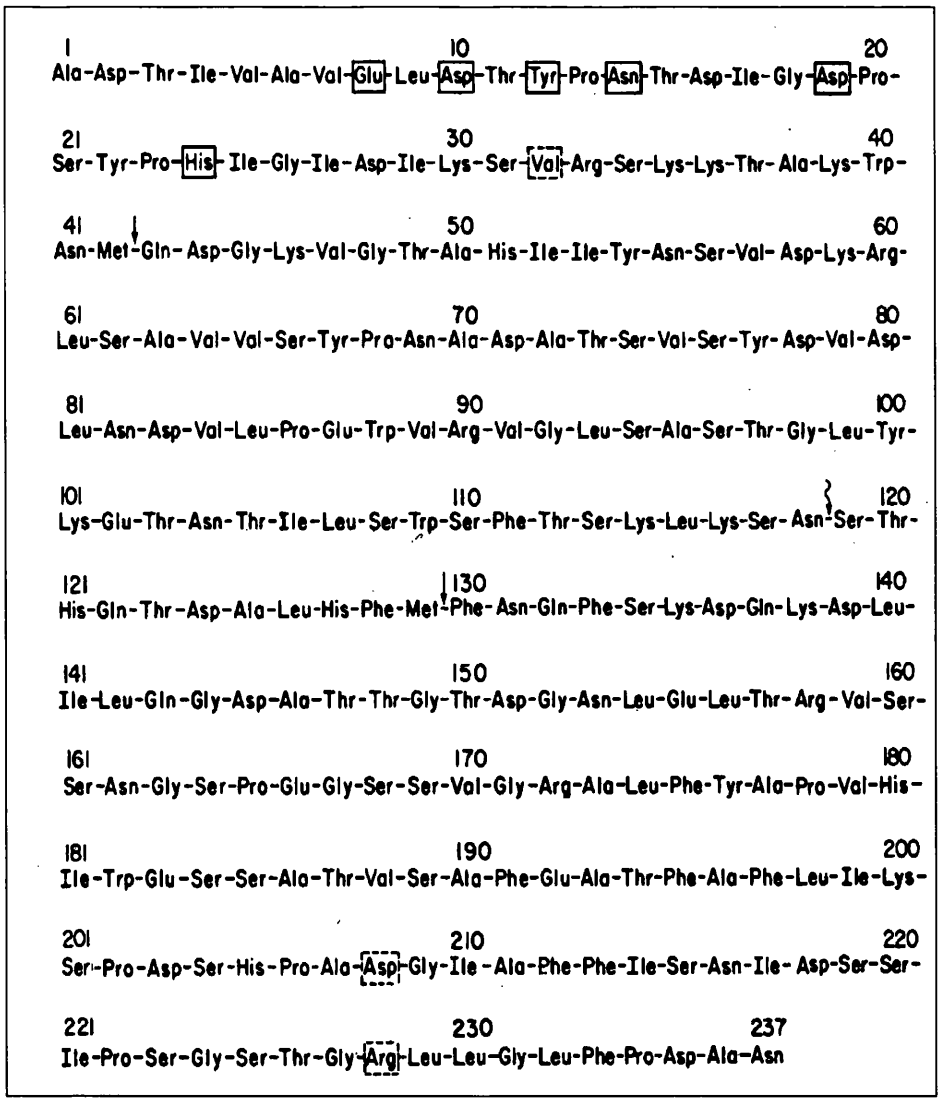


Figure 1.4 Amino acid sequence of mature Con A

Residues interacting directly (solid lines) or through water molecules (dashed lined) with the metal ions of Con A are enclosed in boxes. The sites of cleavage by CNBr are indicated by solid arrows and the natural cleavage occurring in some molecules of Con A is showed by a wavy arrow. Reproduced from **(Becker et al., 1976)** The Molecular Structure of Concanavalin A. In: **Bittiger, H., and Schnebli, H. P. (Eds) Concanavalin A as a tool,** Wiley and Sons pp 33-5

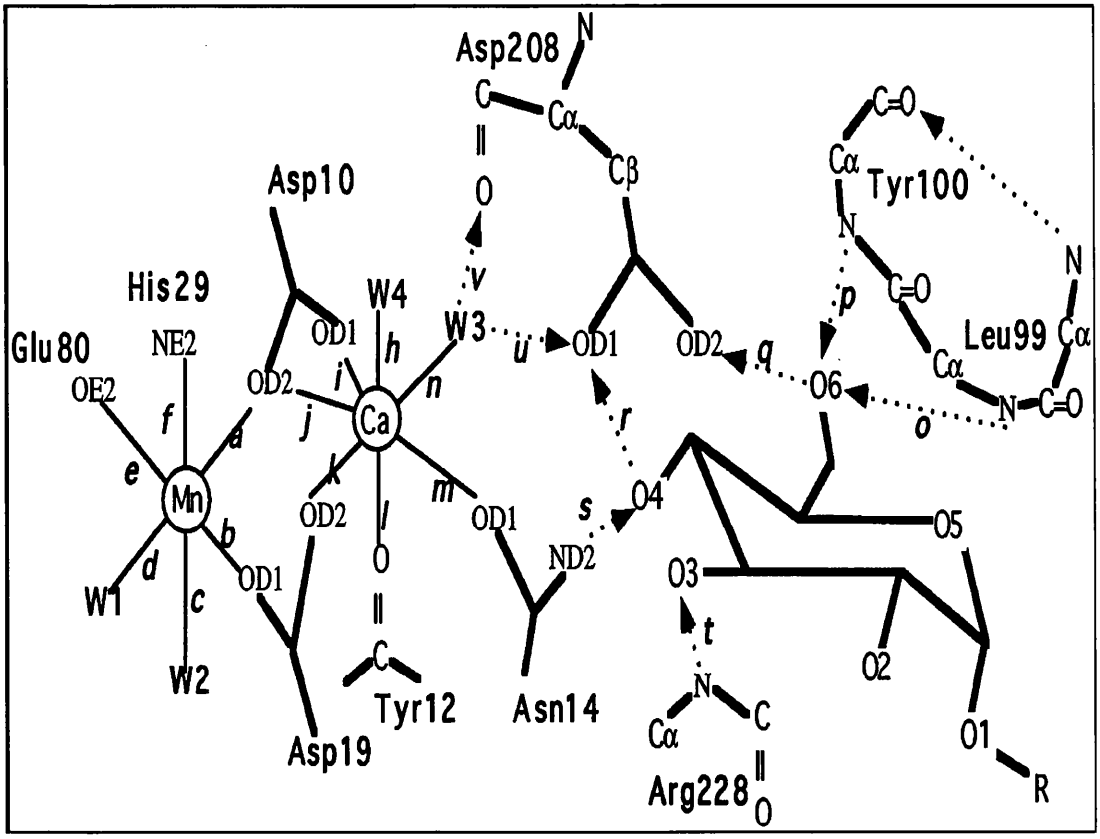


Figure 1.5a Metal–Ligand and saccharide hydrogen bonding distances for Con A subunits

Metal–Ligand Distances and Saccharide Hydrogen Bonding Distances for the Four Subunits. Reproduced from (**Hamodrakas et al., 1997**). Journal of structural biology **118:** 23–30

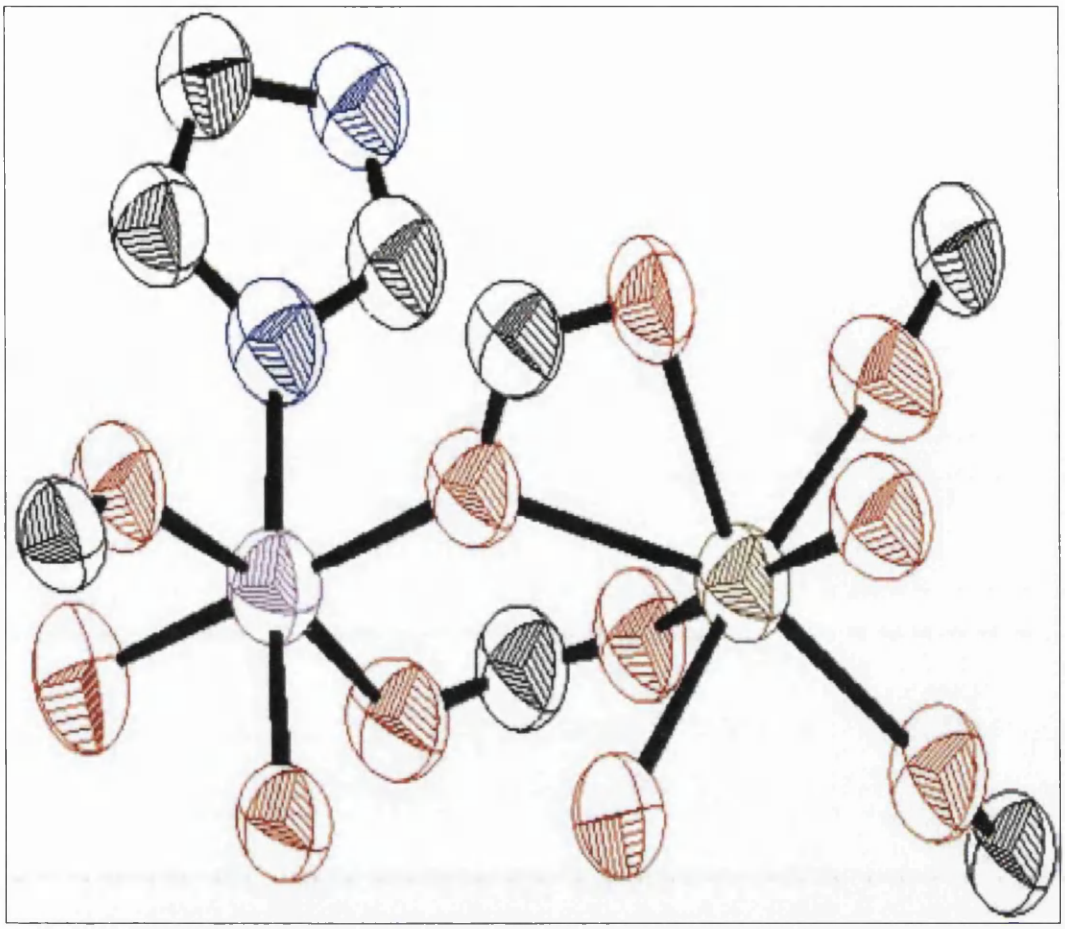


Figure 1.5b Metal-binding sites of Con A

The metal-binding sites S1 and S2 depicted in terms of atomic displacement ellipsoids. Reproduced from **(Deacon, A., 1997).** J. Chem. Soc., Faraday Trans., 93(24).

reported by (Srinivas et al., 2001) that oligomerisation may victimise the lectin's from acquiring it's full biological activity. Dimeric Con A was found to possess binding affinities for sacchrides containing 1 and 2 or 5 mannose sugars like these of the tetrameric structure. Figure (1.9) indicates the three structures of Con A (monomer, dimmer, and tetramer).

1.5.3 Covalent and Three Dimensional Structures

The amino-acid sequence using dansyl-Edman procedure contains 237 residues and three-dimensional structures of the lectin Con A has been determined. The sequences of three short stretches were assigned on the basis of x-ray crystallographic data. Interpretation of an electron density map at 2-Å resolution indicates that the predominant structural element is an extended polypeptide chain arranged in two anti-parallel pleated sheets or ß- structures. Residues not included in the ß -structures are arranged in regions of random coil. One of the pleated sheets contributes extensively to the interactions among the monomers to form both dimers and tetramers. The positions at which Mn2+, Ca2+, and saccharide are bound to the protein, and the point of cleavage for the formation of the naturally occurring fragments A_1 and A_2, have been tentatively assigned. Both metal-binding sites are at least 20-Å removed from the position at which saccharides are bound. The saccharide-binding site is a deep pocket of approximately 6Å × 7.5k × 18k, the inner portion of which is occupied by hydrophobic residues (Edelman et al., 1972).

1.5.4 Binding and functional properties

The binding of Con A to various structures via hydrophobic interactions has been studied using a variety of physicochemical assays. It is found that Con A binds to non-polar compounds such as the plant auxin /Iindoleacetic acid and its structural analogue tryptophan and that this binding is independent of the saccharide binding activity normally associated with the lectin. The results of equilibrium dialysis experiments on the binding of /α-indoleacetic acid were consistent with the presence of a single weak binding site per subunit of protein, having an association constant of about 7 × IO' M-~'. Competition experiments using various non-polar compounds such as oiodobenzoic

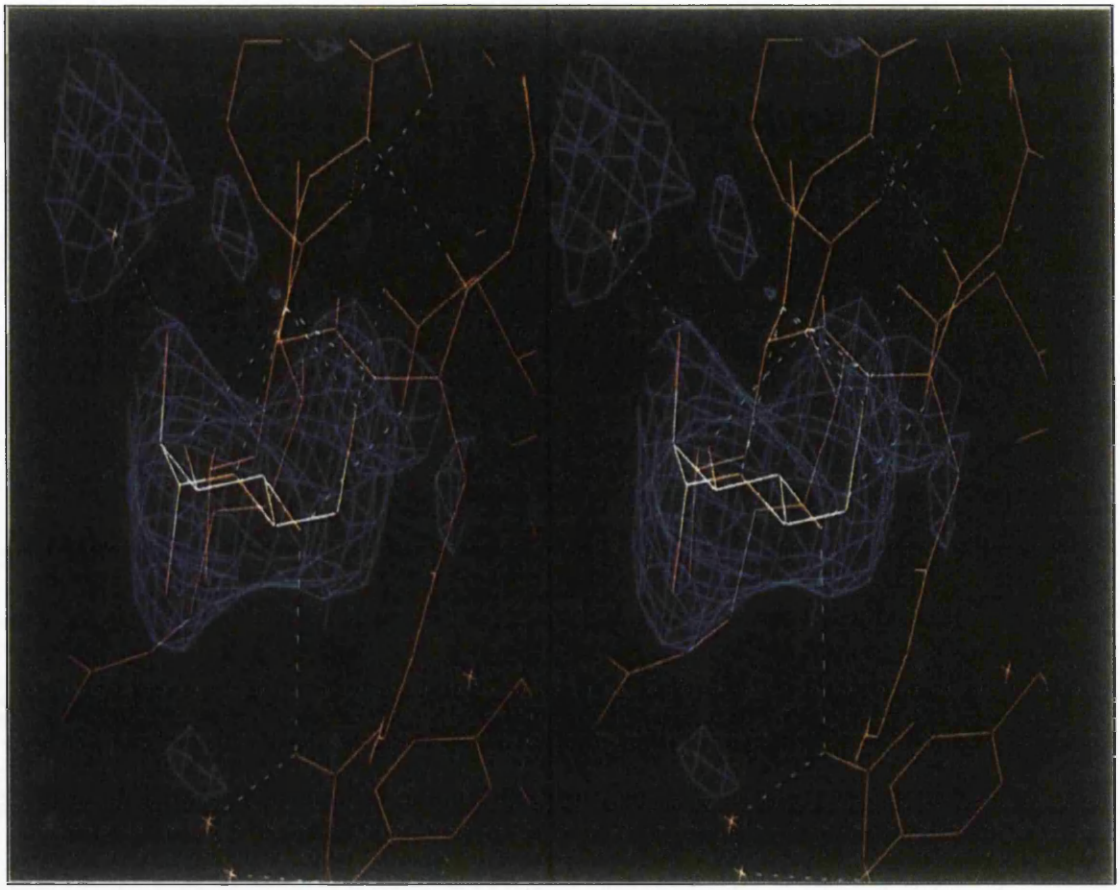


Figure 1.6a Stereoscopic view of the saccharide-binding site

The bound saccharide molecule is superimposed on the difference-electron density map (blue) together with the portion of the protein which constitutes the site. Hydrogen bonds are shown as broken lines. See Figure1.5b for a detailed view of the hydrogen-bonding network. The methyl substitute of 0-1 has not electron density and is not shown in the figure. Reproduced from **(Derewenda, Z. et al., 1989).** The EMBO Journal **8**: no.8 pp.2189-2193.

Figure 1.6b Stereoscopic view of the saccharide-binding site

Showing the network of hydrogen bonds which stabilizes the interaction between the protein and the saccharide. Calcium and transition metal are shown as white and pink spheres. Bonds between the metal ions and their ligands are not shown in this figure for the sake of clarity. Reproduced from **(Derewenda, Z. et al., 1989)**. The EMBO Journal 8: no.8 pp.2189-2193.

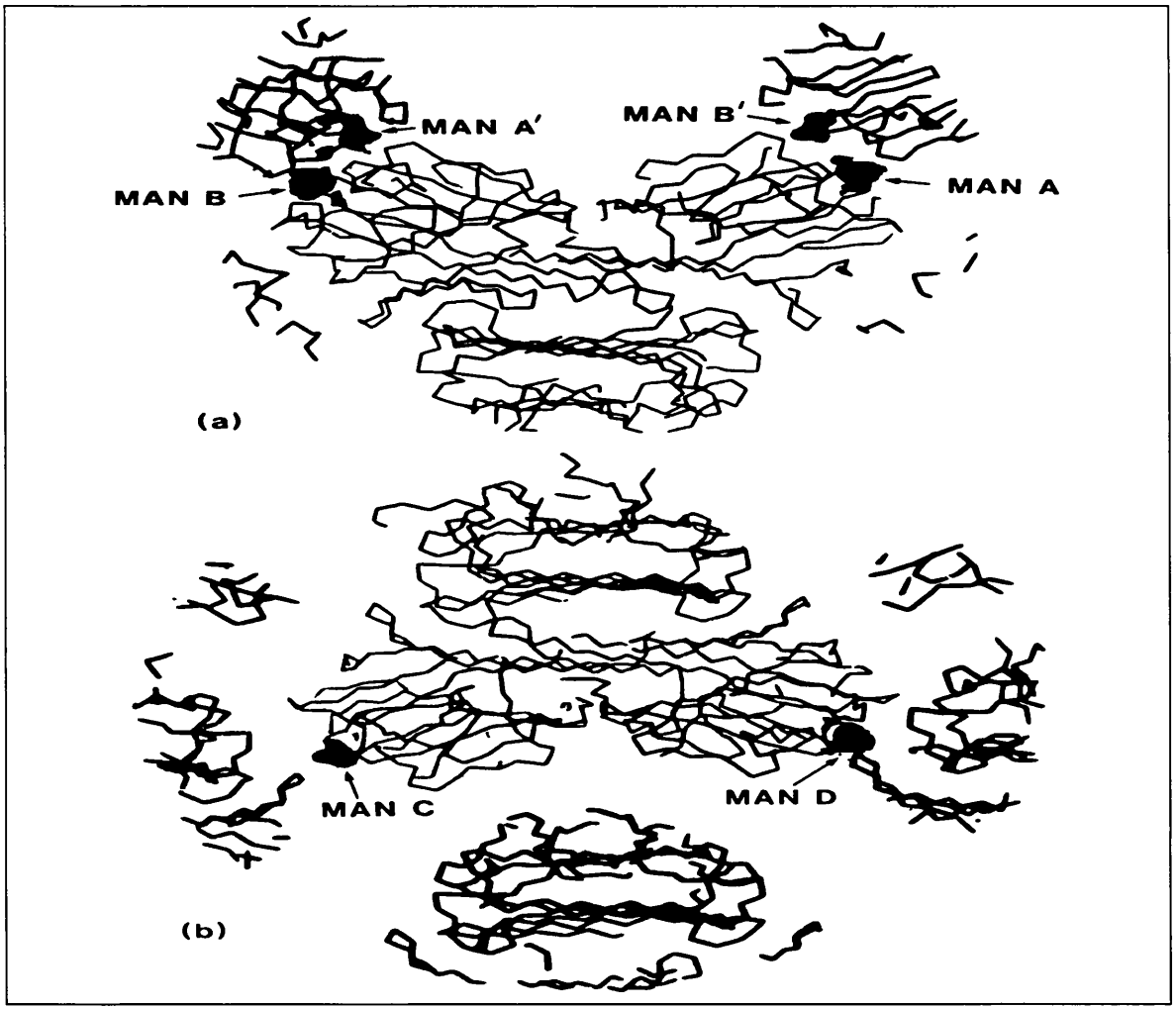


Figure 1.7 The c-carbon backbone of a tetramer of the Concanavalin A

View in the y-direction showing the proximity of the bound mannoside molecules in one Concanavalin A dimer to the mannoside bound to the adjacent tetramers. Each of the two pairs of proximate sites is related by a local pseudo-dyad axis normal to the xz-plane. (b) View in the x-direction showing the proximity of each of the bound mannoside molecules in the second Concanavalin A dimer to the surface of an adjacent tetramer. Locations of bound mannoside are labeled MAN. Reproduced from **(Derewenda, Z. et al., 1989).** The EMBO Journal 8: no.8 pp.2189-2193.

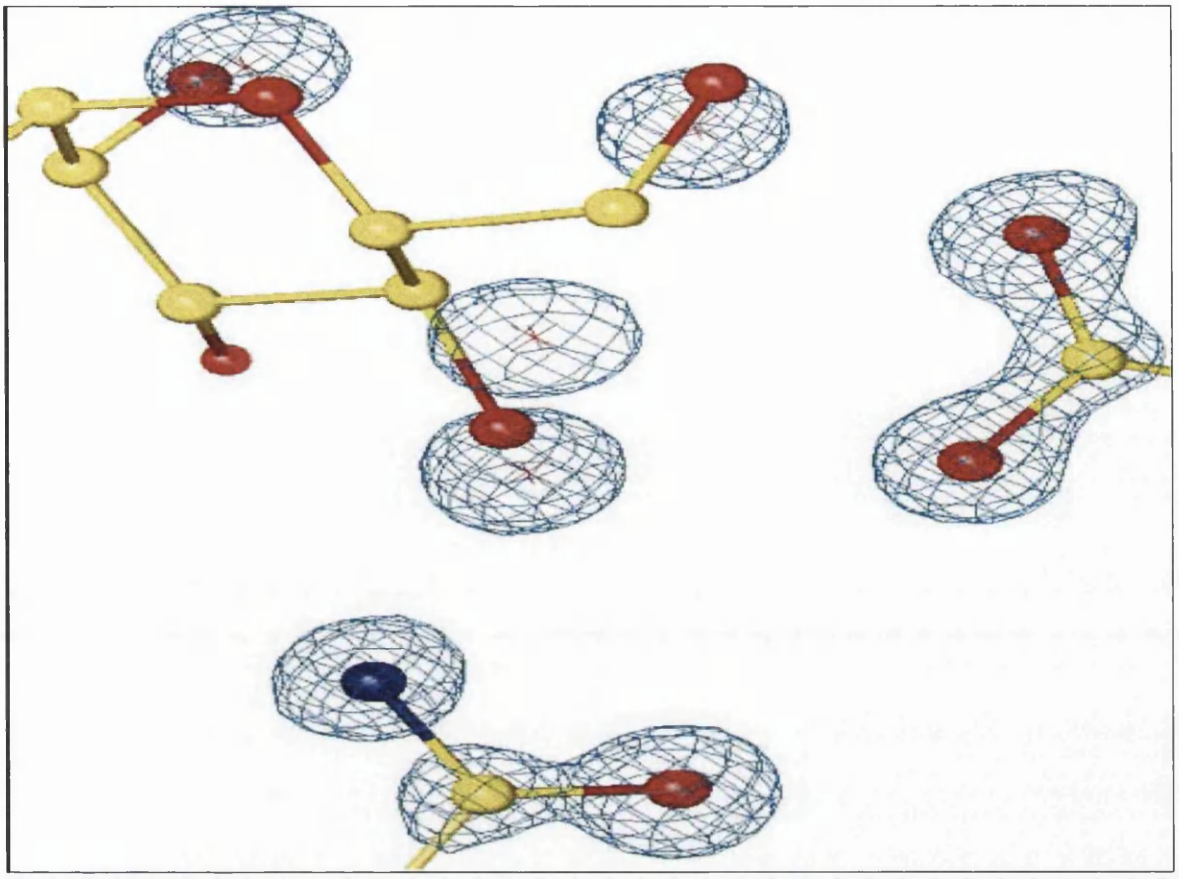


Figure 1.8 Methyl a-D-mannopyranoside (PDB 5CNA)
(Overlaid onto the saccharide-binding site of saccharide-free Concanavalin A at 0.94 Å resolution)

Electron density map (2F indicates the positions o [Fc) of water molecules (top left), Asp208 (top right) and Asn14 (bottom left). Reproduced from **(Deacon, A., 1997)**. J. Chem. Soc., Faraday Trans., 93(24).

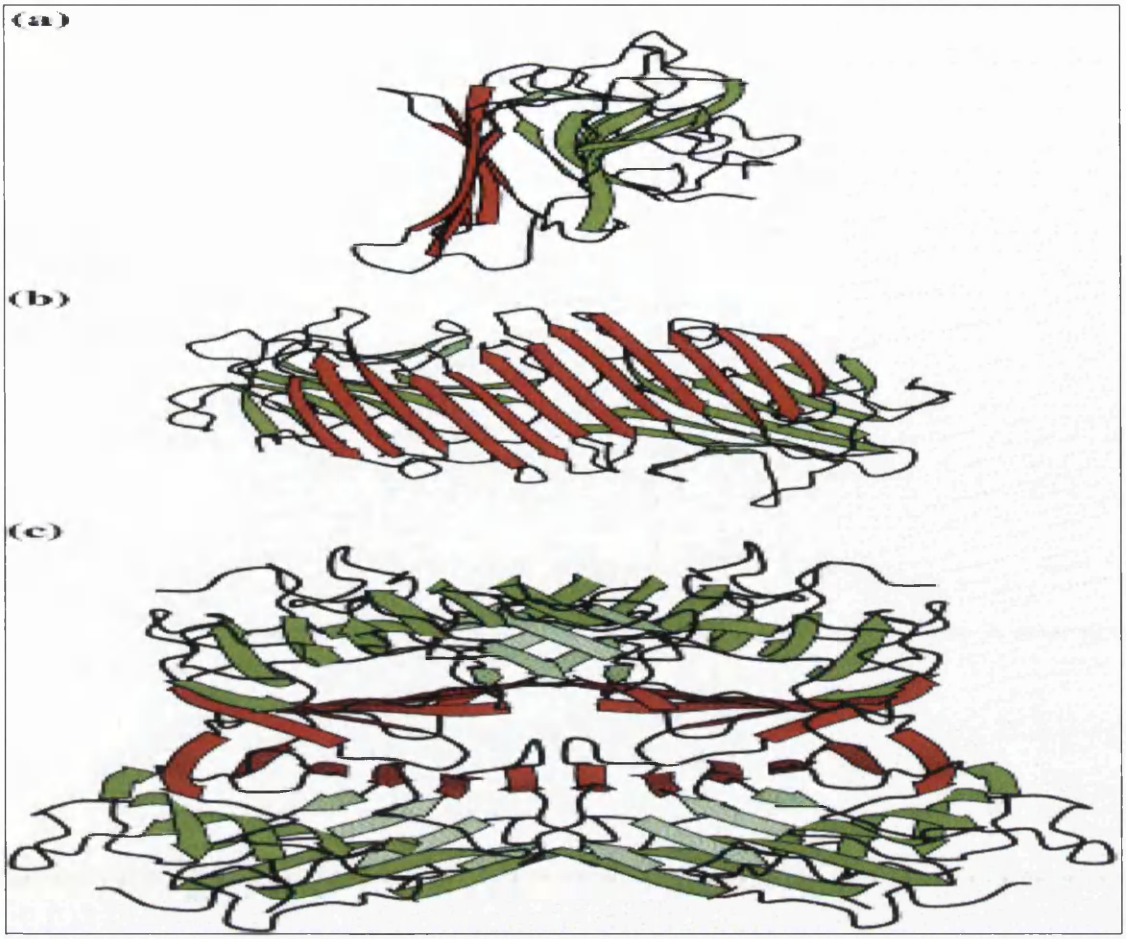


Figure 1.9 Tertiary structure of Concanavalin A (monomer, dimer, and tetramer)

(a) Tertiary structure of Concanavalin A monomer described as 'Jelly roll', (b) dimerisation in Con A (c) the tetramerisation of the two dimers. Adapted from **(Srinivas, V. R., Bhanuprakash Reddy, G., Ahmad, N., Swaminathan, C. P., Mitra N., and Surolia A., 2001)**. Legume lectin family, the 'natural mutants of quaternary state', provide insights into the relationship between protein stability and oligomerisation. Biochim. Biophys. Acta. 1527: 102-111).

acid suggested that this hydrophobic binding site is located in the same cavity which binds the iodine containing ligand as shown by x-ray crystallography.

Con A also binds to lipid vesicles composed of dipalmitoylphosphatidylrholine . This binding to lipid membranes raises the possibility that the synergistic effects of Con A and tetradecanoyl phorbol acetate on lymphocyte mitogenesis may be due in part to an interaction between the lectin and the phorbol ester.

1.5.5 Biosynthesis of Con A

Con A is synthesized in the cotyledons of developing jack been seeds and accumulates in the protein bodied of storage-parenchyma cells. Thereby, Con A is synthesized as an inactive glycoprotein precursor termed (glyco-pro-Con A). Different pathways are followed to produce the peptide backbone of mature form of Con A like deglycosylation, endoproteolytic cleavages, and ligation of peptides. Mature Con A is not a glycoprotein although its initial precursor (pro-Con A) is glycosylated. The precursor pro-Con A, after being secreted to the endoplasmic reticulum and transported to the protein bodies is activated by deglycosylation to pro-Con A (Bowles et al., 1986; Chrispeels et al., 1986) this is a vital stage in the lectin activation Figure 1.10 (Jones (1995). Moreover, (Sheldon et al. (1992 and 1998); Ramis et al. (2001) indicated that an acidic medium is needed beside the deglycosylation to activate the pro-Con A in the plant.

Deglycosylation occurs after the peptide splicing on the 15-residue loop (on the loop of the residues 120-135) of the pro-ConA approximately in the middle of the molecule (Figure1.10). Figure 1.11: has shown the cleavage and ligation of Con A biosynthesis indicating two intermediates comprised of A chain of apparent Mr 18 K and composed of 9 residues of the C-terminal extension of pro-Con A and 4 residues from the spacer peptide corresponds to residues 1-118 of mature Con A. Furthermore, (B chains) of apparent Mr 14 K precursor's equivalent to residues 119-237 of mature Con A (Figure1.11b). Figure1.11d shows ligation at residues 118-119 resulting in rearrangement of sections A and B of apparent Mr 30 K when Mature Con A sequence is synthesized in the jack beans. No junction at 118-119 present in the mature plant produced two native fragments, one a peptide of apparent Mr of16 K (A chain)

equivalent to residues1-118 the other of apparent Mr of 14 K corresponding to residues 119-237 (B chain), (Figure1.11e), (Min and Jones, 1994).

1.5.6 Purification of Con A

The present project is dealing with the refolding part of the purification of Con A affinity chromatography and dextrans as Sephadex G-75 are used. Here in this section is a brief description of the technique which has been firstly mentioned by Sumner (1919) who indicated the isolation of the crystalline form by dialysis of sodium chloride extract of the jack bean meal. Although, the method is described as a simple technique that uses aqueous acetone extract of jack bean meals, methods have been developed to purify Con A that widely differ from the first time. Scientists like Goldstein and Poretz (1986) developed a method exploiting the sacchride binding characteristics of the lectin in affinity chromatography that made the purification of crude Con A accessible. Columns of "Sephadex" (G-25, G-50, and G-75) are prepared and equilibrated with 1.0M NaCl (as the method prefer high salt buffer) in addition to metal ions like (Mn^{2+} and Ca^{2+}).Though, Sephadex G-75 and G-50 are perfect for affinity chromatography, Sephadex-25 is not recommended due to cross-linking to use as a matrix for Con A. Glucose or other sugars in low concentrations are used to elute Con A from the dextran. The method produces yields of 2-3grams of Con A per 100gram of jack bean meal.

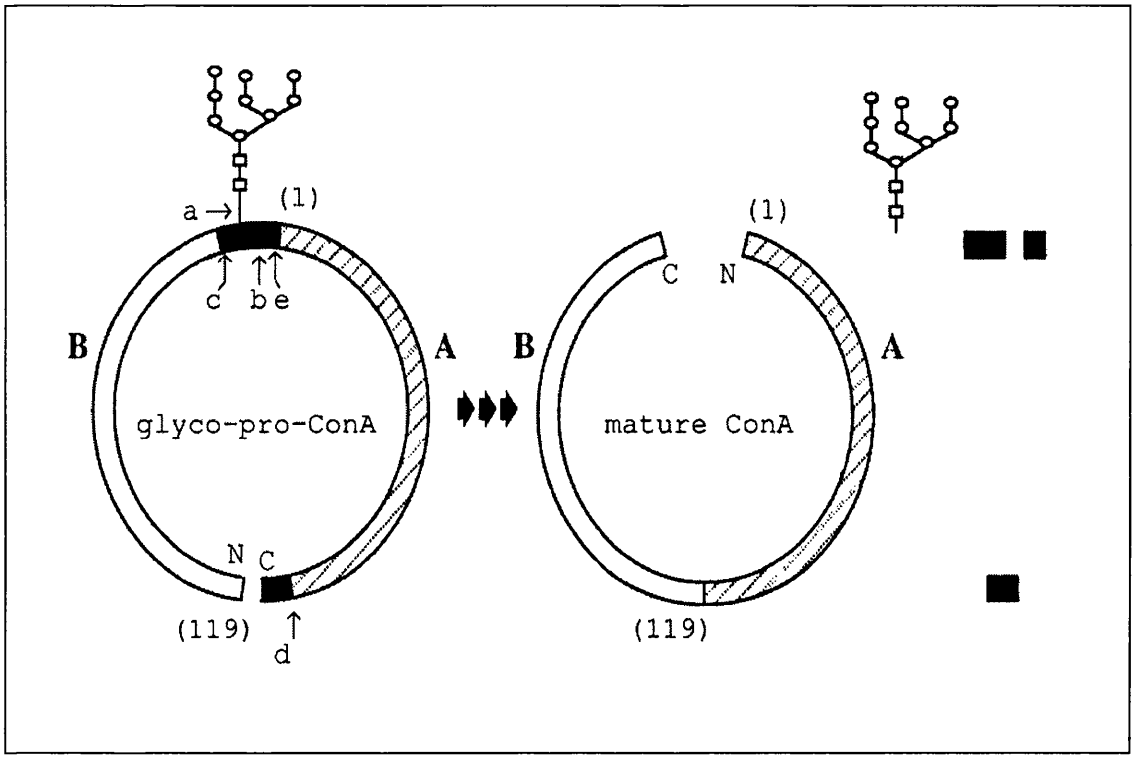


Figure 1.10 Post-translational modifications during Con A biosynthesis

Summary of processing events converting glycosylated pro-Con A to mature Con A during development of immature jack bean cotyledons. Six intermediates, including stages where two polypeptide chains are Ribbon representation of the Con A tetramer (crystal structure) e distinct, have been omitted. Amino- and carboxy-termini are indicated by N and C, and the numbers in brackets are residue positions of mature Con A. The 9 residue C-terminal extension and the 15 residue intervening sequence are shown as solid infill [■]: the latter carries one N-linked high-mannose oligosaccharide (○, Man; □, GlcNAc). During processing in the plant inactive glyco-pro-Con A is first de-glycosylated (arrow a) - resulting in appearance of lectin activity. An endopeptidase then cleaves (arrows b, c, d, e) on the C-side of four asparagine residues, and residues (118) (arrow d) and (119) are ligated enzymatically. Splicing thus results in a transposition in the linear arrangement of the protein sections designated B [□] and A [▒], with multi-staged excision of the glycosylated spacer sequence and nonapeptide tail. The mature lectin becomes a circular permutation of most of the protein sequence of its own precursor. Adapted from: **Jones, D. H. (1995)** "Folding, activation and protein splicing of recombinant Concanavalin A precursors: an exceptional protein to prove some rules" Chap. 20 *In* Perspectives on Protein Engineering and Complementary Technologies, Geisow, M. and Epton, R. (eds.) Mayflower Worldwide, Birmingham (ISBN 0-9515735-2-7).

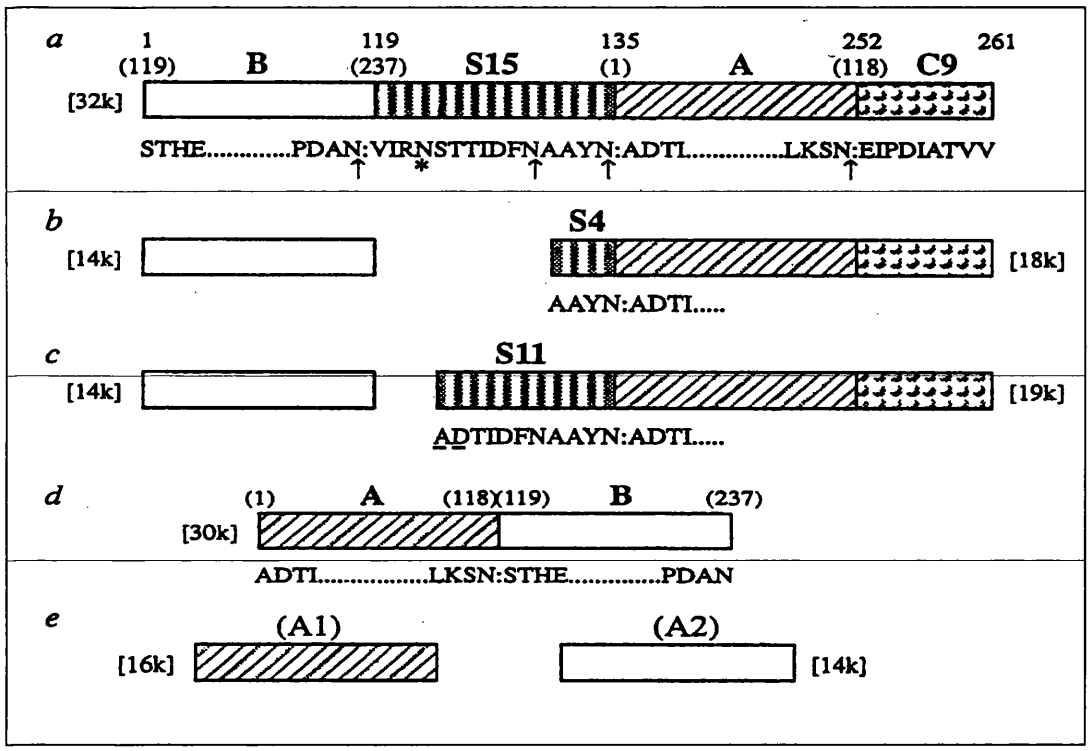


Figure 1.11 Cleavages and ligation generating Con A
(in jack beans and comparison with recombinant proteins used to reconstitute splicing in vitro)

(a) Pro-Con A showing section B, intervening 15 residue sequence (S15), section A, C-terminal 9 residue extension (C9). (Regions A and B not drawn to scale). Residue numbers are shown above with the numbering corresponding to mature Con A in round brackets. Apparent M_r, (based on electrophoretic migration) is shown in square brackets. Amino acid sequence is given below in single letter code with boundaries between regions shown (:), glycosylation site shown (*) and known cleavage sites following asparagine residues indicated by arrows. (b) Processing intermediates cleaved from pro-Con A in jack bean detected by pulse-chase labelling (not all intermediate stages are shown). (c) Recombinant precursor A and B chains constructed from pIN-pro, expressed separately in E .coli, refolded together and affinity purified. The region derived from the intervening sequence is truncated to 11 residues (S11) the first two of which are altered (underlined) as a result of DNA manipulations. (d) Mature Con A sequence generated in developing jack beans by ligation at residues 118-119 resulting in rearrangement of sections A and B. (e) Native fragments detectable in mature seeds, originally designated A1 and A2. (Reproduced from **Min, W. and Jones, D.H., (1994)**. Nature Struct. Biol., 1: 502-504,).

1.6 Previous Work in This Laboratory:

Optimising the refolding conditions of a recombinant protein, produced as an insoluble inclusion body by over expression in *Escherichia coli*, may explain certain facts about the mechanism of how proteins fold (section 1.2). Many problems face the large scale productions of correct refolding process in industry are represented in low yield, high cost of reagents and equipment (section 1.4). On the other hand, laboratory refolding procedures differ in the use of reagents and vessels considered unsafe and costly in industry as detergents, chaotropes, extreme pH or temperature. The use of model proteins as tools to unveil different sights about the refolding mechanisms is widely accepted in scientific research, however, the way a protein can refold cannot be generalised as proteins vary in how they refold to their native structures. More than fifteen years have already passed since Min (1992) first used the pIN series of expression systems and recommended host strains in our laboratory. Great works of successors accomplished vast improvements in yield and protein quality. Mei Li (2001) work on expression and refolding of ConA variants added a mark in identifying the different aspects of expression and purification for this model. Thus, the present project continues from where they have stopped in improving the protein yield and quality through following a thread to optimise the refolding of this model protein. The knowledge gained from literature (section1.5) indicated the problems that are challenging the correct refolding of Con A and here in this project focused on the prevention of aggregation by altering some of the environmental factors that aid in the refolding process Figure (1.12).

1.7 Aims of the Project

The objective of this project is to increase the yield and the quality of *in vitro* refolded proteins over-expressed in *E. coli* as insoluble inclusion bodies. This can be achieved through tackling the problems that initiate aggregation in refolding media (Figure 1.12). The specific aims of the project are:

1- To investigate the use of glycerol as a refolding aid that reduces hydrophobic interactions, which are the main cause of aggregation and thus low yield.

2- To evaluate the use of low concentrations of guanidine hydrochloride to decrease aggregation during refolding.

3- To optimising the dilution factors and dilution protocols to improve the overall protein yield.

4- To investigate Triton X-100 as an aggregation suppressor.

5- To explore more deeply how the refolding of Con A proceeds using Triton X-100 as a tool to determine the time that the denatured protein takes to reach the completely folded structure.

6- To compare the effect of several sets of temperature ranges from as low as -5.5^{0}C to as high as 37^{0}C in order to adjust the optimum *in vitro* temperature at which Con A properly folds to its native structure

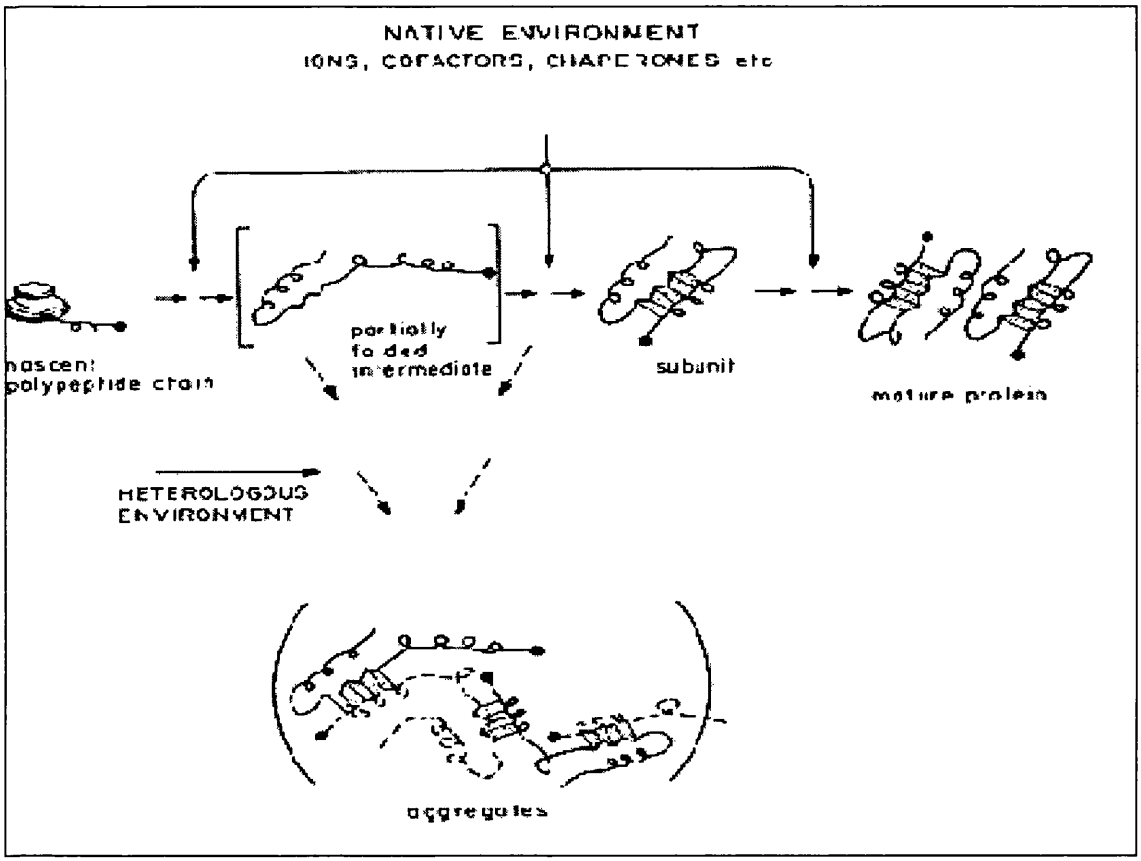


Figure 1.12 In vivo hypothetical folding and maturation pathway for a dimeric protein

The result of the pathway depends on the presence of the proper environmental conditions, including cofactors, chaperones etc. A heterologous environment can influence the passage of partially folded intermediates towards the aggregated state. (Reproduced from **Mitraki, A., and King, J., 1989).** Protein folding intermediates and inclusion body formation *Biotechnology* 7: 690-697)

Chapter Two

Materials and Methods

2.1 Materials Used In Experiments

2.1.1 Chemicals and Reagents

Unless other-wise mentioned, all chemicals and reagents used were of analytical grade. These reagents were used without any further purification.

2.1.2 Bacterial strains

Strain	Genotype
BL21 (DE3)	F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dcm* (DE3)

2.1.3 Plasmid Vectors

The parent expression vector used was pET-26b (+). The plasmid map of the expression vector is shown (Figure 2.1). The Con A variant used is:

pET 0MAT_ Mature-Con A coding sequence with *ompA* ribosome binding site and secretion signal inserted as *Xba*I – *Bam*HI fragment into corresponding corresponding sites of pET-26b(+).

2.2 DNA Manipulations

2.2.1 Growing colonies

The bacterial cells BL21 (DE3) purchased from Novagen were plated out on LB plates without antibiotics and incubated at 37^0C for 24 hours. Single colonies were used fresh next day.

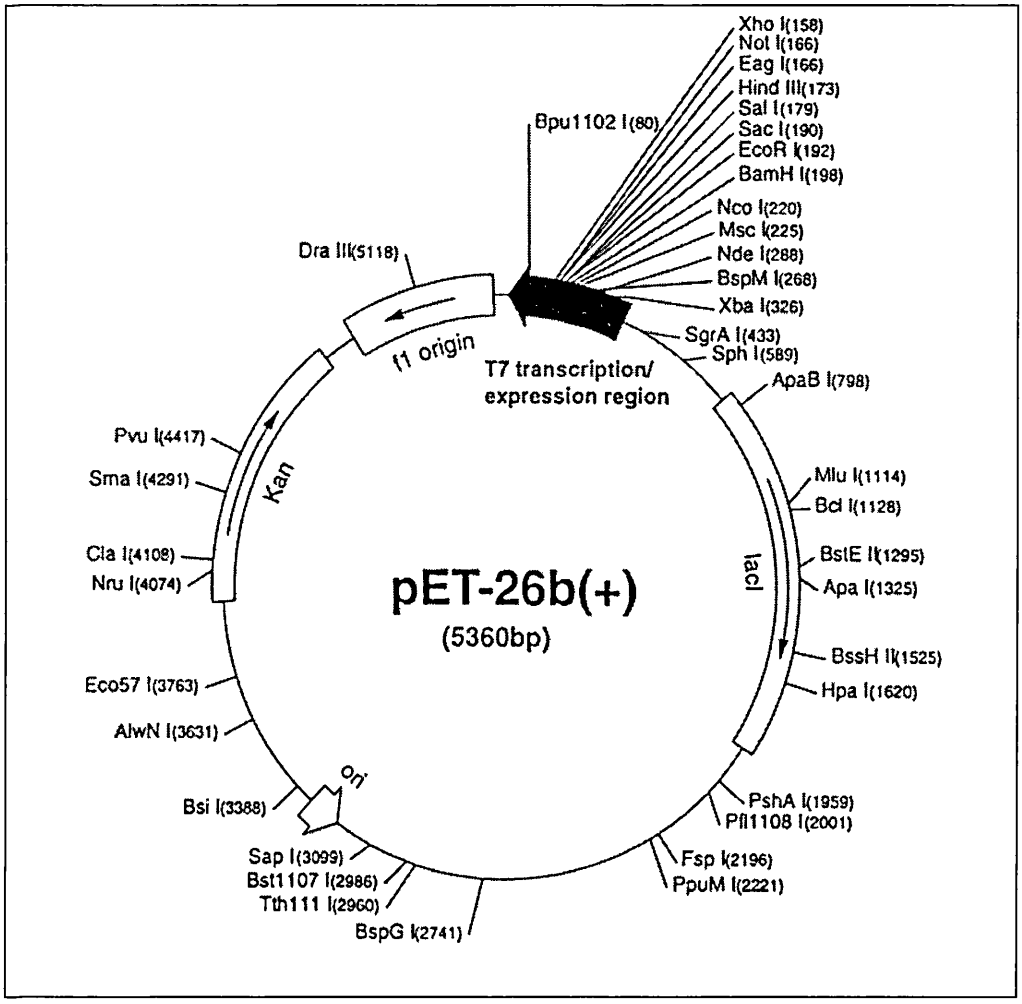


Figure 2.1 Plasmid Map of Expression Vector pET-26b (+) [Novagen]

Unique restriction enzyme sites only are shown. Features: **ori** – plasmid origin of replication (for ds DNA), **Kan** – Kanamycin resistance gene, **f1 origin** – origin of replication for ss DNA in presence of helper phage, **lacI** – gene for *lac I* repressor protein. The **T7 transcription/expression region** includes a T7promoter/lac operator site (before *Xba*I site), and a T7 transcription terminator (after *Bpu*1102 I site). [Figure adapted from Novagen Technical Bulletin #71 (1993).]

2.2.2 Preparation of electro- competent cells

100 ml LB (section 2.4.1.1.) was inoculated with a single colony of *E. coli* strain BL21 DE3. The culture was incubated at 37^0C from mid afternoon until the next day. 5 ml of the incubated culture were used to inoculate 500 ml of LB medium in a 2, 000 ml flask (warmed overnight). The cells were incubated at 37^0C and their growth was monitored at 600 nm until the absorbance reached 1.5. At this density the cells were chilled on ice for 20 minutes. For all the subsequent steps of washing the cells, tubes and solutions were kept ice-cold. The cells were divided into two 250 ml aliquots and were harvested by centrifugation at 4,000X g for 15 minutes using a Beckman Avanti J 20 rotor. The pellets were re-suspended each in 250 ml of sterile 10% glycerol (pre-chilled on ice) making slurry of the pellet in a small amount of the 10 % glycerol first and spun at the same speeds. Again the pellet was discarded and the cells were re-suspended in 125 ml of 10% glycerol using the slurry method. The re-suspended cells were combined into one centrifuge bottle, balanced using discarded glycerol from first wash, and spun again. Subsequently the supernatant was poured off and the pellet was re-suspended in 20 ml of sterile 10% glycerol. Balancing against 20 ml 10% glycerol followed centrifugation at 4,000 x g for 15 minutes at 4^0C. Again the supernatant was discarded and at this step the pellet was re-suspended in 2 ml sterile 10% glycerol. Finally 0.5 ml aliquots of the cells were placed in 40 pre-cooled labelled 1.5 ml micro centrifuge tubes. The tubes were sealed, snap frozen with liquid nitrogen and stored at -70^0C.

2.3 Transformation of competent cells by electro-poration

Both DNA sample and the competent cells from -75^0C were placed on ice. 0.5µl of DNA (plasmid preparation) was transferred to the tubes containing 50 µl of competent cells. After 20 minutes on ice, approximately 55µl of DNA and cells solution was put in a pre-chilled electro-poration cuvette. On the electro-poration device (Biorad Micropulser^{TM}) the bacteria programme was selected. The cuvette was tapped down to ensure that the cells were between the plates. It was then inserted into carrier. The carrier was pushed in until the electro-porator clicked home firm. The electro-porator was activated and the pulse cycle was delivered by pushing the charged button for a few seconds until a short beep was heard. Afterwards the cuvette was removed from the

sample chamber and immediately the cells were recovered with 1 ml of LB and glucose 0.5% (w/v), and transferred back to the culture tube for 1 hour incubation at 37^0C.
Both 200µl and 20µl aliquots of transformed cells were plated respectively onto surface-dry LB agar plates containing Kanamycin (30 µg/ml) and incubated overnight at 37^0C (the 20µl of aliquot of electro-poration mixture was mixed with 180µl LB and spread).

2.3.1 Production of starter inoculum

In order to get more reliable growth and expression , frozen inoculum must be used, so next day, for each plasmid trans-formant five typical colonies were selected and placed into five universals containing 7.5 ml LB, 0.5 % (w/v) glucose and Kanamycin (30 µg/ml). The colonies were incubated at 37 ^{0}C and their growth was monitored at 600 nm until the absorbance had reached 1.7. At this density 750 µl of sterile 80% (v/v) glycerol (final concentration 8%) was added and the universals were swirled to mix. The samples were rapidly placed on ice to arrest growth. For each of the 5 typical colonies of each plasmid trans-formant 9 aliquots of 0.6 ml and one of 1 ml chilled culture was placed in pre-cooled 1.5 ml micro-centrifuge tubes. The tubes were sealed and snap frozen in liquid nitrogen. Finally, they were stored in a plastic storage box in the -80^0C freezer.

2.4 Protein Techniques

2.4.1 Preparation of Solutions and Media

2.4.1.1 Preparation of LB and Kanamycin liquid medium (2 litre)

Eight flasks containing 500 ml of LB medium each and 100 ml in a Schott bottle were prepared by mixing 42g of tryptone (LUCHEFA), 21 g of yeast extract (Oxoid) and 42 g of NaCl (Fisher Scientific) in water up to 4.1 litre. The solution (original pH about 6.9) was adjusted to pH 6.0 with 5 M HCl. The medium was autoclaved at 121^0C for 20 min and subsequently was left to cool on the bench. Just before the media were used, kanamycin was added to each flask and the bottle to a final concentration 30 µg/ml (0.5 ml stock per 500 ml) using aseptic techniques.

2.4.1.2 Preparation of Stock Kanamycin Solution

Kanamycin was dissolved to a final concentration of 30 mg/ml in water. The solution was sterilised by 0.22μm filtration and afterwards 0.5 ml aliquots were aseptically pipetted into sterile microfuge tubes and stored at -20^{0}C.

2.4.1.3 Preparation of IPTG

2.38 g of IPTG (MW=238.3) were dissolved in 20 ml of distilled water. The solution was sterilized by filtration through a 0.22 μm disposal filter. IPTG solution was stored in 1 ml aliquots (0.5 M) and stored at -20^{0}C.

2.4.1.4 Preparation of Borate wash buffer (1 litre)

20 mM Borax [7.63 g Na2B4O7.10H2O (Di-sodium tetra borate 10H_2O-Analar] and 5 mM EDTA [1.86 g Na_2EDTA.2H_2O (Diaminoethanetetra-acetic acid disodium salt-DUCHEFA] were added to 1 distilled litre water. The resulting pH was 9.1-9.2, no further adjustments were required.

2.4.1.5 Preparation of MOPS metals buffer (1 litre)

1 M NaCl [58.44 g (Fisher Scientific)] and 50 mM MOPS [10.5 g (Sigma)] were made up to 1 litre with distilled water. The resulting solution was adjusted to pH 7.0 using 10 M NaOH. Afterwards, 1 ml was added of each of 1 M $CaCl_2$ and 1 M $MnCl_2$ giving a final concentration of each metal of 1 mM. $CaCl_2$ and $MnCl_2$ were not added before adjusting pH to avoid precipitation. Finally, 0.1g (0.1 g/L) of NaN_3 (Sodium Azide SLR-Fisons) was added to the solution.

2.4.1.6 Preparation of MOPS Refolding Buffer

The preparation of MOPS refolding buffer is the same as the preparation of MOPS metals buffer except the final concentration of calcium and manganese were 10 mM, therefore 10 ml of each of the stock of 1 M metal solutions Ca^{2+} and Mn^{2+} were added to 1 L of the buffer.

2.4.1.7 Preparation of Metals Stock Solutions

1 M Calcium Chloride Dihydrate, 14.70 g were dissolved in 100 ml of distilled water in a graduated volumetric flask. For 1 M Manganese Chloride Tetrahydrate, 19.79 g were dissolved in 100 ml of distilled water in a graduated volumetric flask.

2.4.1.8 Preparation of 8 M Guanidine-Hcl for Protein Purification

152.85 g of guanidine-HCl (Sigma) (8 M) were at once dissolved by adding 80 ml of distilled water to take the solution to 200 ml exactly. The solution was then placed in a 50^0C water bath to dissolve and the volume was checked again when cooled at room temperature. Finally, it was stored at room temperature.

2.4.1.9Preparation of Sephadex G-75

The solution is prepared in advance by weighing 5 g of Sephadex G-75 powder (40-120 μm dry bead diameters, Pharmacia), 100 ml of MOPS metals buffer were added to the powder in a glass Duran bottle. Then this was placed in a boiling water bath to swell the Sephadex gel. The water was brought to boiling on a hot plate and the solution was simmered for half-hour. Subsequently the solution was left on the bench to cool overnight. Then was stored at 4^0C.

2.4.1.10 Preparation of Mannoside Elution Buffer

0.1942 g (10 mM) Methyl-α, D-Mannopyranoside (Sigma) was dissolved in 100 ml of MOPS metals buffer and a trace of Azide was added.

2.4.2 Growth and Induction of Cells

As described previously (Section 2.4.1.1) eight conical flasks (4 litre), each including 500 ml of sterile LB medium and a bottle (Schott) containing 100 ml of LB medium were prepared and pre-wormed overnight at 37^0C. 100 µl of Kanamycin solution was added to 100 ml LB medium in the bottle. 5 ml of the starter medium (LB and Kanamycin) was transferred from the bottle to eight universal tubes (Sterilin). Also, 0.5 ml of thawed kanamycin stock solution was added to the 500 ml LB media in the conical flasks, followed by 1 ml of 1 % (v/v) Antifoam (Sigma) sterile suspension.

Eight tubes containing 0.6 ml of frozen inoculum were thawed in the 37 ^{0}C water-bath for 30 seconds and vortexed for 1 second. Then, using the hood and a sterile pipette, 0.5 ml of the thawed inoculum was added to each of the eight universals containing the starter medium. The top of the tubes were loosened half a turn and secured with a piece of sticky tape. The universals and the conical flaks were then incubated for 2 hours at 37^0C/250 rpm. After 2 hours the starter culture in the universals was added to the eight conical flasks and incubation was continued.

After about an hour the absorbance of the culture at 600 nm (using a WPA Light-wave 2000 spectrophotometer) was measured at hourly intervals by taking 1 ml sample with an automatic pipettor sterilised with alcohol and using distilled water as a blank. At the point where (usually after 3.5 hours) the absorbance reading at 600 nm was <2.0, 1 ml of IPTG solution was added to each of the eight flasks to a final concentration of 1mM and growth was continued for a further 3 hours.

After the three hours, a final cell density reading was measured and then mixing of the cells was conducted for the eight flasks to remove any variations that might occurred in the overall growth culture, then (first 250 ml of the 500 ml culture) were harvested by centrifugation (Beckman J2-21, rotor JA10) at 10,000 rpm for 20 minutes at 4^0C. The supernatant was discarded and the above repeated for the other 250 ml of the culture in the same centrifuge bottle. Finally the cell pellets were placed at -80^0C.

During the growth and induction of cells aliquots were taken from each of the eight flasks for pH Measurements to follow, changes in pH during growth. Samples were taken (1) Before inoculation, (2) Before IPTG addition, (3) and at an hourly interval to harvest.

2.4.3 Extraction of Cells (Preparation of Insoluble Material)

The pellets from 500 ml culture volume were each thawed in 40 ml borate wash buffer containing 1 % (v/v) Triton X-100. Then the pellets were re-suspended as much as possible with a big glass rod and then in a shaker incubator for 10 min. The solution (40 ml approximately) was then divided between two Oakridge style centrifuge tubes (50 ml size) and sonicated (3x1 min) on hard packed ice to break the cells open. Maximum power (tuner) with the probe about 5 mm from the bottom of the tube was used. The cell lysate was incubated at room temperature for 10 min to 1 hour with gentle mixing on the rocker. The cells debris obtained was balanced and centrifuged at 20,000 rpm for 20min at 4^0C in a JA-20 rotor (Beckman J2-21). Then the supernatant was discarded and the pellet was kept. The tubes were drained by standing up side down on a piece of tissue. The drained pellets were then re-suspended in 20 ml borate wash buffer without Triton using 1 X 1 min sonication. The solution was mixed at room temperature for 10 min to half hour and then was spun down at 20,000 rpm for 20 min at 4^0C in a JA-20 rotor (Beckman J2-21). Again the supernatant was poured off and the pellets were now re-suspended in 20 ml of Milli Q H_2O using 1 X 1 min sonication and then spun down at 20,000 rpm for 20 min at 4^0C. Finally, washed pellets were frozen at -80^0C.

2.4.4 Extraction of Proteins

The frozen pellets were thawed for a few minutes at room temperature on the bench. A glass rod was used to break up the pellets as finely as possible and they were dissolved in 3 ml of 8 M Guanidine HCl dissolved in MOPS refolding buffer. Mixing of extracts was also conducted at this point for every set of experiments to remove any variations that could arise before either at the growth level or the Guanidine –HCl level and to insure exact refolding conditions. Afterwards, the equal amount of extract for 2 dissolved pellets was poured (for example 2 x 3 ml = 6 ml) into one Oakridge tube.

2.4.5 Refolding

The refolding techniques differ with each study but in general it was as follows: The 6 ml guanidine extract containing the solubilised protein was refolded by rapidly adding it to 180 ml (30X dilutions) of ice cold MOPS refolding buffer while swirling. The

refolded extract was kept on ice for 1 hour and then on the bench for another 1 hour. Then the extracts were put into two JA-14 (250 ml) tubes and were centrifuged at 14,000 rpm for 30 minutes at 4^0C. Finally, the supernatant was filtered through a tea strainer to cold receiving flasks and the clear supernatant was then loaded onto a column of sephadex G-75 overnight.

2.4.6 Purification of Con A Affinity Chromatography

For the purification of Con A, affinity chromatography was used. It was performed in three steps that are explained through sections 2.4.6.1 to 2.4.6.3 below. Generally, the steps involved (1) the column preparation, where the column was packed with about a 5 ml bed of sephadex G-75 equilibrated in MOPS metals buffer, (2) The loading of the refolding extract to the column overnight at 0.1 ml/min and washing of the column with about 50 ml MOPS metal buffer at 0.5 ml/min and (3) the elution of the Con A variants from the column with 12 ml of 10 mM Methyl-α, D-Mannopyranoside in MOPS metals buffer at 0.2 ml per minute. Generally, four columns were run in parallel at one time.

2.4.6.1Column preparation

First, it was checked that the column was connected to the reservoir (allowing inflow when removing the plunger). Then the plunger was removed from the column. The taps on outgoing lines were turned to intermediate point to plug. The incoming line tubing was clamped in pump, to stop dripping. The column was emptied using a Pasteur pipette and washed three times with water by vortexing to disperse previously used Sephadex. 5-8 ml of Sephadex slurry was then added to the column topped up with MOPS metals buffer and was allowed to settle for 5 min. Subsequently the taps were turned to connect lines to empty flasks for outflow. The column was topped up again with MOPS metals buffer to overflowing and the plunger was inserted. The plunger was stopped 1 cm above the Sephadex bed, its level was checked and the plunger screw was tightened. The incoming line was switched back to the MOPS reservoir. The plugs were removed from the outgoing tubes. Finally, to compact the column buffer was pumped at 1ml per minute (Pump Setting 2 @ X 10), for 2 minutes.

2.4.6.2 Sample Loading and Washing Away of Contaminants

As soon as it was ensured that the entire flow path was connected to the sample, 180 ml of refolded extract was pumped down the column of Sephadex G-75 at a flow rate of 0.1 ml per minute (Pump Setting 2 @ X1). After all the Con A protein would have bound to the Sephadex, other contaminants were washed away with 50 ml MOPS metal buffer at 0.5 ml per minute (Pump Setting 10 @ X1).

2.4.6.3 Elution

The chart recorder and A_{280} monitor [flow cell spectrophotometer (Pharmacia)] was turned on and the range was set at to 2.0 A_{280}. The chart speed was selected to be 0.5 mm/min, units were adjusted to mm/min and 10 mV full-scale spans were selected. The A_{280} monitor was turned to "SHORT" and zero to zero was adjusted on the chart recorder. Also the flow cell spectrophotometer was turned back to 2.0 and the base line was adjusted to 2 %. The pen was put down and the chart feed was turned on. The protein was eluted from the column at 0.2 ml/min (Pump Setting 4 @ X1) using 12 ml 10 mM Methyl-α, D-Mannopyranoside in MOPS metals buffer. During elution 12 fractions (1 ml each) were collected. When elution was finished, a post elution system wash was performed. The columns were washed with 50 ml of MOPS metal buffer from reservoir at 0.5 ml/min (Pump Setting 10 @ X 1). Finally when washing was finished, the flow cell spectrophotometer, chart recorder and pumps were switched off and the tubes plugged and released from the pumps pressure plates.

2.4.6.4 General Design of Experiments

Different protocols of the refolding and binding of the lectin Con A stages were investigated in the current project. Some studies explore the effects of additives like (glycerol, extra guanidine-HCl, and TritonX-100) on the purification of Con A. Therefore a specific design was depicted in Figure 2.2; indicates the practical plan followed in these studies.

2.4.6.5 Spectrophotometric Analysis

Absorbance readings at 280 nm were taken for each of the 12 fractions. A Beckman DU 650 Spectrophotometer was used and conversion factor 0.877 was selected to convert A_{280} readings to mg/ml [$E_{280}^{1\%}$=11.4 for Con A]. 1 ml of Mannoside elution buffer as used above was used as a blank. Elution fractions from each flask with absorbance readings greater than 0.1 were pooled in a small centrifuge tube. The volume in each tube was measured with a small glass cylinder and used for further yield calculations. The absorbance of each tube of pooled fractions was taken three times at 280 nm. 800 µl of Mannoside was used as a blank and for each measurement 200 µl of sample was then mixed with the 800 µl of Mannoside elution buffer to give a fivefold dilution. This ensured that the readings are within the most accurate range of the spectrophotometer. The average of three close readings was used to calculate concentration of purified protein for each flask. Finally, a wave-length scan at 250-300 nm was performed and the absorbance at 260 nm and 280 nm were read to calculate the A_{280}/A_{260} ratio which if > 170 indicates the less contaminants like (nucleic acids, lipid, and carbohydrate molecules) and the purity more of the refolded protein.

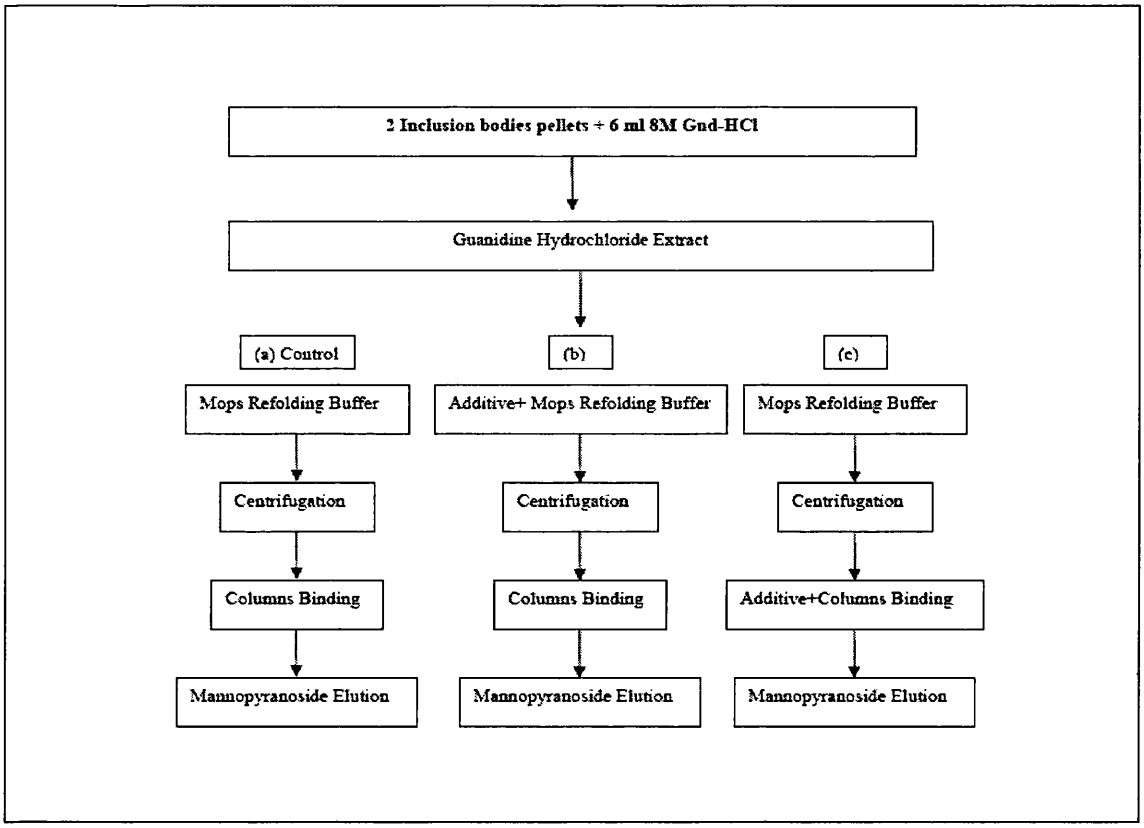


Figure 2.2 General schemes for an additive

Shows the practical steps used (a) Control is at 0.27 M Gnd-HCl concentration i.e. with no additions just the guanidine hydrochloride extract used at start to denature the protein. (b) Additive additions at the refolding stage, and (c) Additive additions before the binding to affinity chromatography columns or the after refolding stage.

2.4.7 Yield Calculations

The final yield is expressed in milligrams per litre of growth culture (mg/l culture), and is calculated as follows:

A280×0.877= mg/ml

Mg/ml × 5 = Mg/5ml

Mg/5ml× mls/tube = mg/tube

Mg/tube×2 = mg/l culture

Where:

A280 = absorbance at wave length 280.

0.877 = conversion factor of absorbance into milligrams per millilitre used.

5 = dilution factor used to ensure that the readings are within the most accurate range of the spectrophotometer.

Mls/tube = volume of the eluted protein fractions with absorbance readings greater than 0.1 pooled in a small centrifuge tube.

Mg/tube = milligrams in 500 millilitre culture

2.4.8 Statistical Definitions Used

Definitions and formulae are given below:

For N observations of X_i values:

Mean (m) = $(\Sigma X_i)/N$

Standard Error of the Mean (SEM) = $S/\sqrt{[N-1]} = \hat{S}/\sqrt{[N]}$

Where:

S = standard deviation of the Mean

$\hat{S}$ = standard deviation of small population size.

Chapter Three

Results

3.1 Introduction

Recombinant proteins over expressed in *E. coli* face many challenges during refolding. Aggregation is the major cause of low yields that different studies have tried to overcome. The present project is trying to improve the refolding renaturation yields by introducing some additives like glycerol, guanidine hydrochloride (Gnd-HCl), and Triton X-100. Moreover, optimise the refolding incubation temperature and dilution protocols. A general scheme of the different experiments used illustrated in Figure 2.2 provides the main idea followed using the different additives. In addition, Figures 3.1 and 3.3 explain the processing scheme for the high-level expression of the protein model Con A.

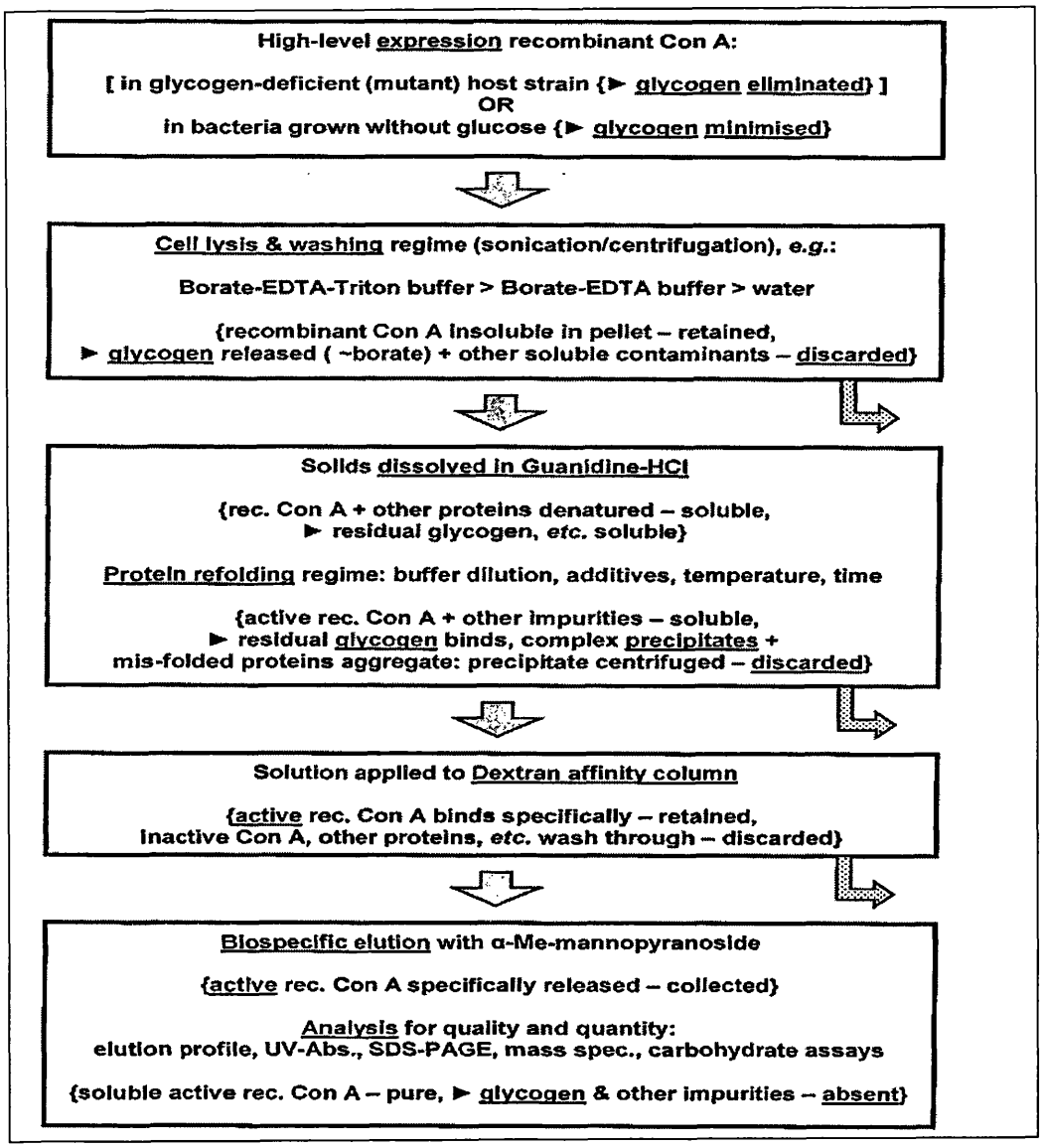


Figure 3.1 Scheme for preferred embodiment method for the purification of Con A

Purification method for a recombinant glucose binding protein. Reproduced with permission from **Jones, H., (2005)**. WO20051987.Patent.

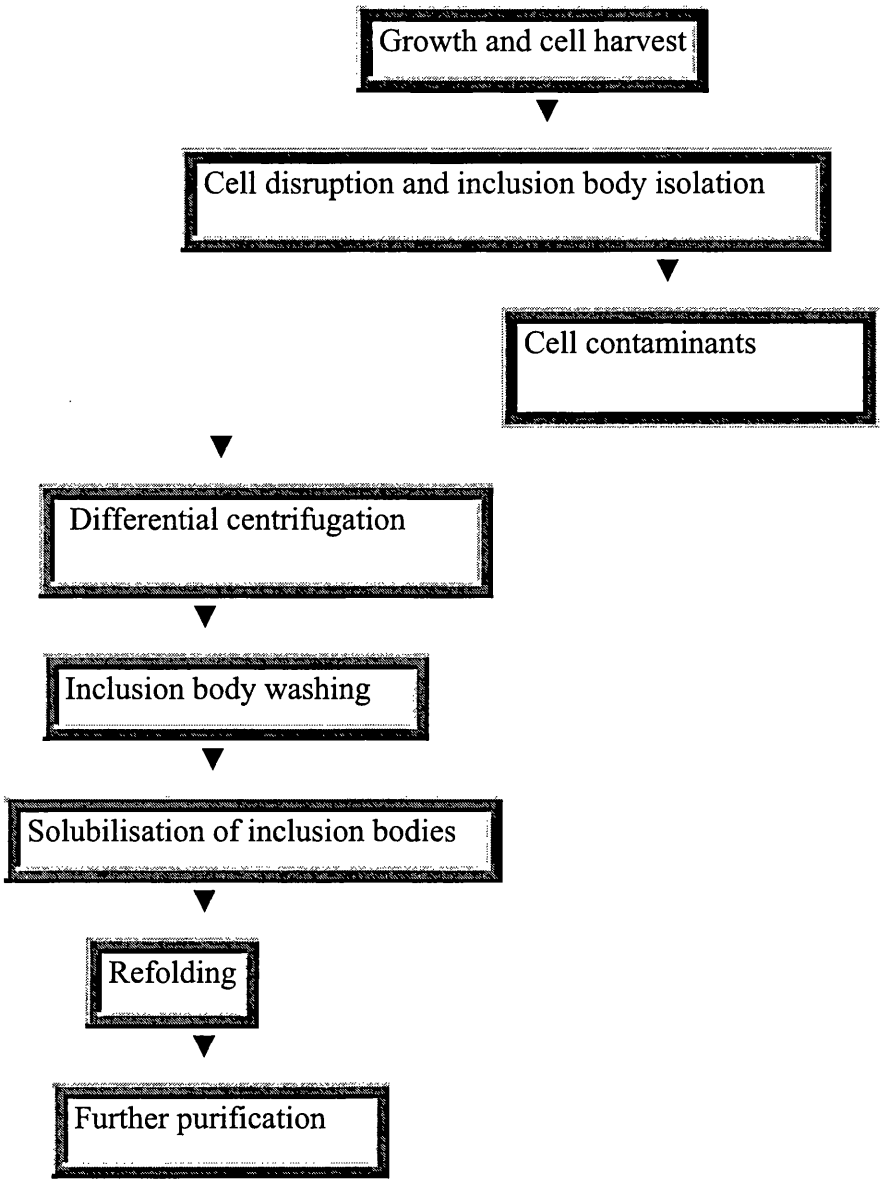


Figure 3.2 The Purification Scheme of Inclusion Bodies In This Laboratory

Illustrating the purification scheme of inclusion bodies over expressed in *e-coli* used in our laboratory.

3.1.1 Glycerol

3.1.1.1 Introduction

The present study targeted protein aggregation as one of the major reasons for yield reduction by introducing glycerol as one of the refolding aids (see section1.4.7.2) to control the effect produced by hydrophobic interactions. It was stated that dilution or dialysis techniques of the refolding of dissolved inclusion body proteins result in aggregated folding intermediates mainly caused by hydrophobic interactions (Fan-Guo et al., 2001). Although purified inclusion bodies provided a tool in the improvement of in vitro refolded protein yield (Rattenholl et al., 2001) and (Nashihara et al., 2000) artificial chaperones (chemical aids) have been introduced to assist in protein refolding (Fan-Guo et al., 2001) and (Kumar et al., 1998).

3.1.1.2 Method

In the current study glycerol was used as a co-solvent additive, at concentrations of 0, 1, 2, 5, and 10 % (v/v) (Figure2.2) to investigate its effects on the folding process. Because this additive will also remain during column loading, its effects on folding alone may be masked by any additional effects on affinity binding. Thus, in order to determine possible direct effects of glycerol on the binding of active protein to the dextran matrix, further experiments were also conducted in which glycerol was added only after the refolding stage. The terms "refolding" or "binding" were determined by the addition of glycerol either at the start of the refolding process or just before loading of the already refolded buffer onto Sephadex G-75 chromatography columns, respectively. So, here as in the general method (refer to section 2.4.5) harvested bacterial inclusion bodies as solid pellets were processed mechanically, washed for optimum purification and solubilised by 8M chaotropic guanidine hydrochloride. Refolding was then carried out by rapidly diluting the 6 ml guanidine denatured extract 30-fold, as aggregate formation is highly protein concentration-dependent, into ice cold MOPS refolding buffer. At this stage the concentration of guanidine hydrochloride normally present is 0.27M. The refolded extract was then kept on ice for 1 hour, at room temperature for another hour then loaded onto a chromatography column of Sephadex G-75 after being subjected to centrifugation. After being loaded for approximately 18 h pure protein was then eluted

with 10 mM Methyl-α, D-Mannopyranoside in MOPS metals buffer. A part of the study used glycerol as an eluant then continued to use the normal elution buffer as previously mentioned. The quality of the purified protein was then confirmed by sharp elution profiles and spectroscopic wavelength scans.

3.1.1.3 Results and Discussion

To block aggregation in the refolding solutions (Table 3.1) indicates the effect of different concentrations of glycerol on the refolding and dextran-binding stages of Con A purification. (a) The refolding stage indicates concomitant reduction in yields with the increase in glycerol concentrations as 10% glycerol scored the lowest yield. (b)Another experiment was conducted using glycerol at 5% and 10% before the binding to the dextran in an attempt to further explore the effect of glycerol on the binding stage indicated decreased amounts of yield with 10% glycerol concentration as the lowest. In a comparison between (a) and (b) parts of (Table 3.1), and (Figure 3.3), it was found that: 5% glycerol concentration indicated a decrease of 29% at the binding stage to the dextran compared to the refolding stage. Moreover, 10% glycerol also showed a decrease of 8.7% at the binding stage compared to the refolding stage (comparison attempted with data having the same control sets). In addition, the elution process was investigated to exploit the effect of glycerol on it by introducing (5% and 10%) as elution buffers instead of methyl- α -D-mannopyranoside, which is the current eluant used (Larson. et al., 1994). However, the yield was insignificant as (Figure.3.4): indicated that glycerol has no significant effect on the elution of the lectin manifested by the small peak of the profile compared to control. However, (Figure.3.5); shows the elution profile and yield when the same sample eluted for the second time with methyl-∝-D-mannopyranoside. Glycerol has produced a negative effect on the refolding stage starting from 1% and to the lowest yield at 10%. Although, scientists like (Meng. et al., 2001) have considered glycerol as one of the most effective additives on the prevention of aggregation and that 10% glycerol was found to be the optimum percentage that completely prevent aggregate formation. The comparison between the two parts of (Table 3.1) and (Figure 3.3); implies that certain amount of purified con A was lost some where between refolding and elution. This means that glycerol doesn't improve the refolding of Con A as expected since it reduces the yield when used during the

refolding stage, and at the same time severely affect the binding of the lectin to the dextran. Before, the work of different research groups (Charron et al., 2002; Derewenda et al., 1989; and Jain et al., 2001a)) indicated (Figure3.6) that glycerol displaces water molecules at certain sites (G1 and G2) of the four subunits of the Con A molecule. However, the interesting part is that glycerol does not just displace water but a recognition phenomenon arises where G1 interacts with participating residues in carbohydrate recognition (Asn14, Leu99, Asp208 and Arg228). Moreover, the glycerol at G4 interacts with Asn44 and Ser201 (peptide recognition), and also involved in the interaction with 15-mer and 10-mer peptides. Therefore, and as a conclusion to the previously mentioned Lopez-Jaramillo, et al., (2004) has argued that: the fact that the G2 site is occupied by a molecule of glycerol in the four monomers of the asymmetric unit support the observations that G2 is playing a vital role in the binding/recognition of ligands. These reports confirm that glycerol affects the binding of the lectin to dextran which agrees with the current study results. Further, the disappointing result of the use of glycerol as an eluant also explained that Con A is more susceptible to methyl-∝-D-mannopyranoside than glycerol. And lead to the assumption that glycerol compete for the carbohydrate binding site with methyl-∝-D-mannopyranoside making Con A less available for elution and, hence, purification. As a conclusion glycerol in this study added an insight to the problem concerning its ability to affect the saccharide binding sites of the lectin.

Glycerol concentration % (v/v) added at stage:	(a) Refolding			(b) Binding		
	Yield (mg/l)	A_{280}/A_{260}	Relative Yield%	Yield (mg/l)	A_{280}/A_{260}	Relative Yield%
0	14.1,16.3,14.6, 13.9 M = 14.8 SEM =0.55	1.8,1.8,1.9, 1.9	100	*14.1,16.3,14.6, 13.9* *M = 14.8* *SEM =0.55*	*1.8,1.8,1.9, 1.9*	*100*
1	15.6,12.8, 10.9 Mean = 13.1 SEM = 1.67	1.9,1.8, 2.0	88.5	Not Determined	----	-----
2	11.2	1.9	75.7	Not Determined	----	----
5	11.6	1.9	78.4	7.3	1.9	49.3
10	5.0, 5.1, 6.6 Mean = 5.6	1.8, 1.9	37.8	4.3	2.0	29.1

Table 3.1 The effect of varying glycerol concentrations on the refolding and dextran-binding stages of Con A purification

The control data (0% glycerol) were also shown in italics for the binding stage. The refolding part of the table indicates a decrease of 11.35, 24.07, 21.36, and 62.24% according to the percentages of glycerol added respectively. Moreover, the after refolding or before binding to the dextran part focuses on the great fall in the yield to 70.85% from control at 10% glycerol.

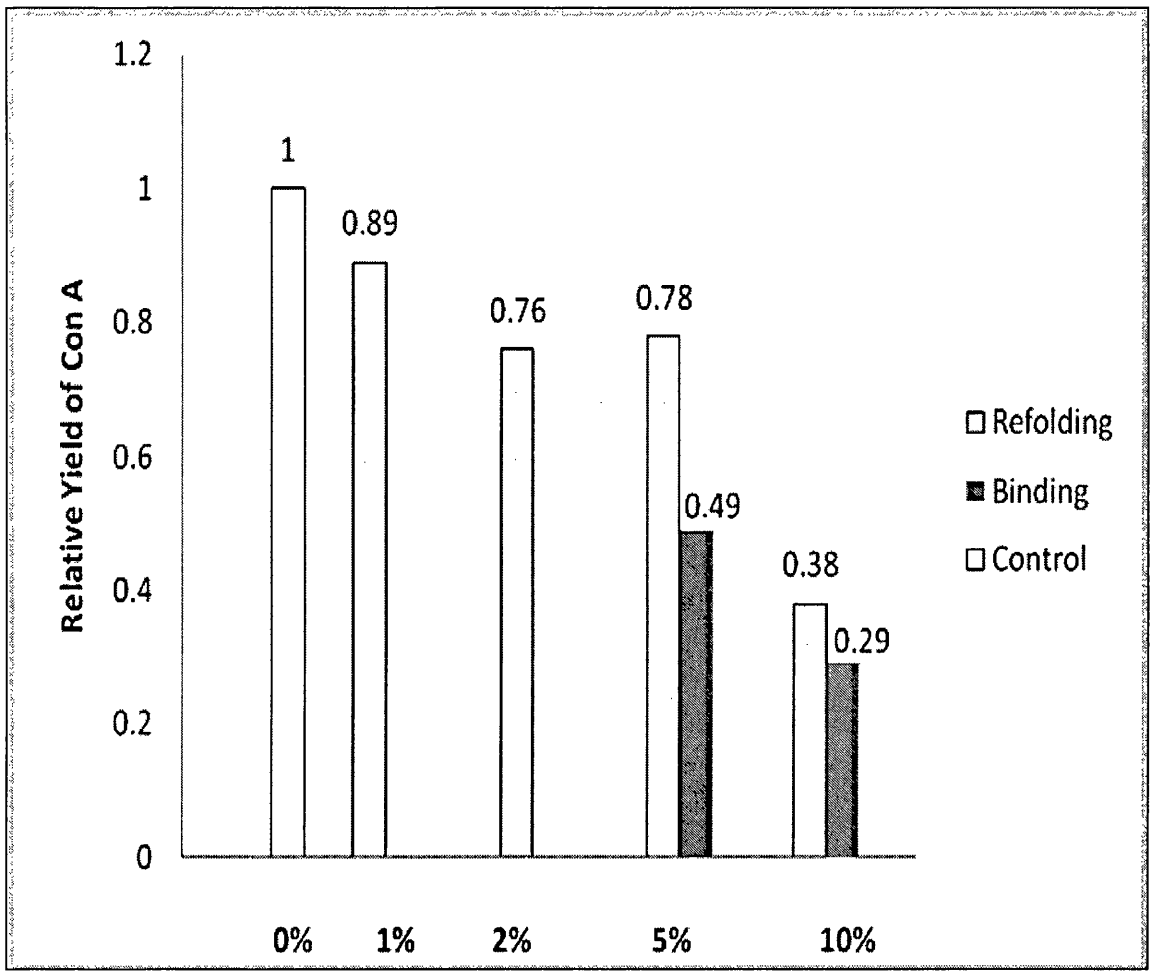
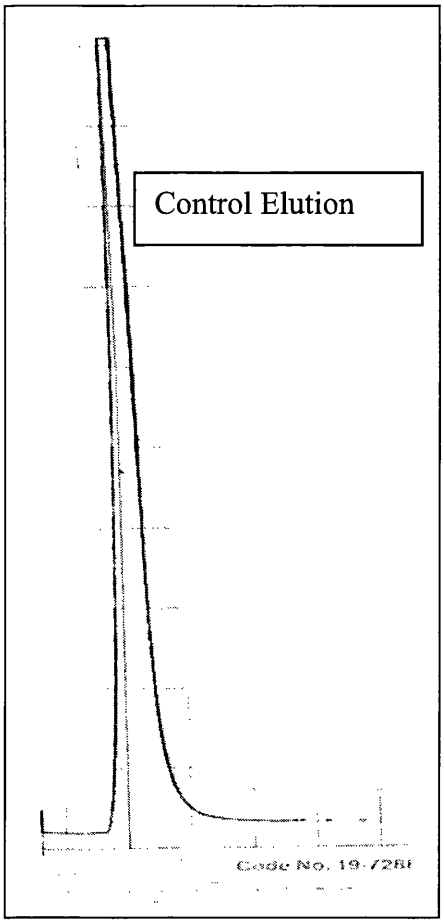


Figure 3.3 Glycerol concentrations (%) added at refolding and binding stages

Shows the relative yield when adding glycerol at 1, 2, 5 and 10 % (v/v) during the refolding stage the after refolding or before binding to the dextran of Con Apurification. Control data (0% glycerol) illustrated (white) in the figure. The decrease in yield is clear at the refolding stages, while it is vividly clear at the binding stage with 10% indicating the lowest.



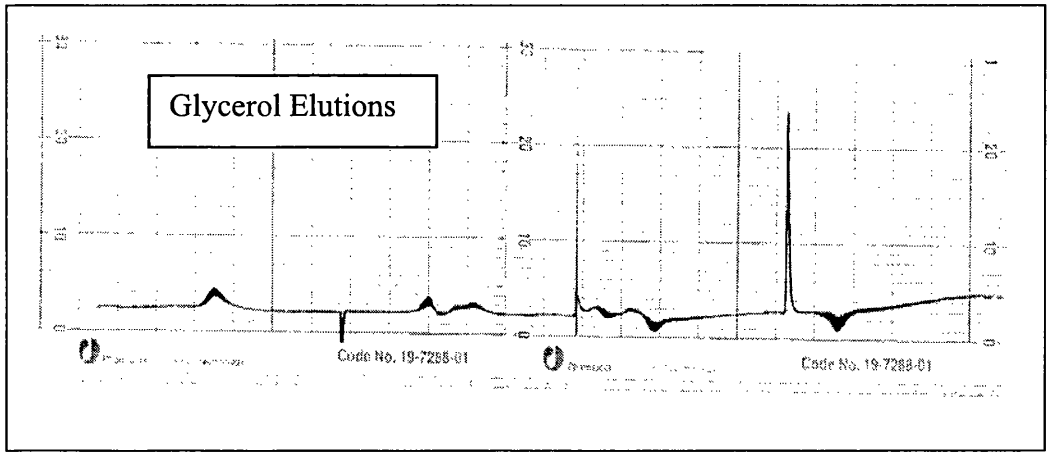


Figure 3.4 Elution profiles using 5% and 10% glycerol concentrations

Shows elution profiles when 5% and 10% glycerol concentrations where used as eluants in place of methyl-∝-D-mannopyranoside the usual eluant. Small peaks indicate no purified proteins were eluted compared to the control shown above.

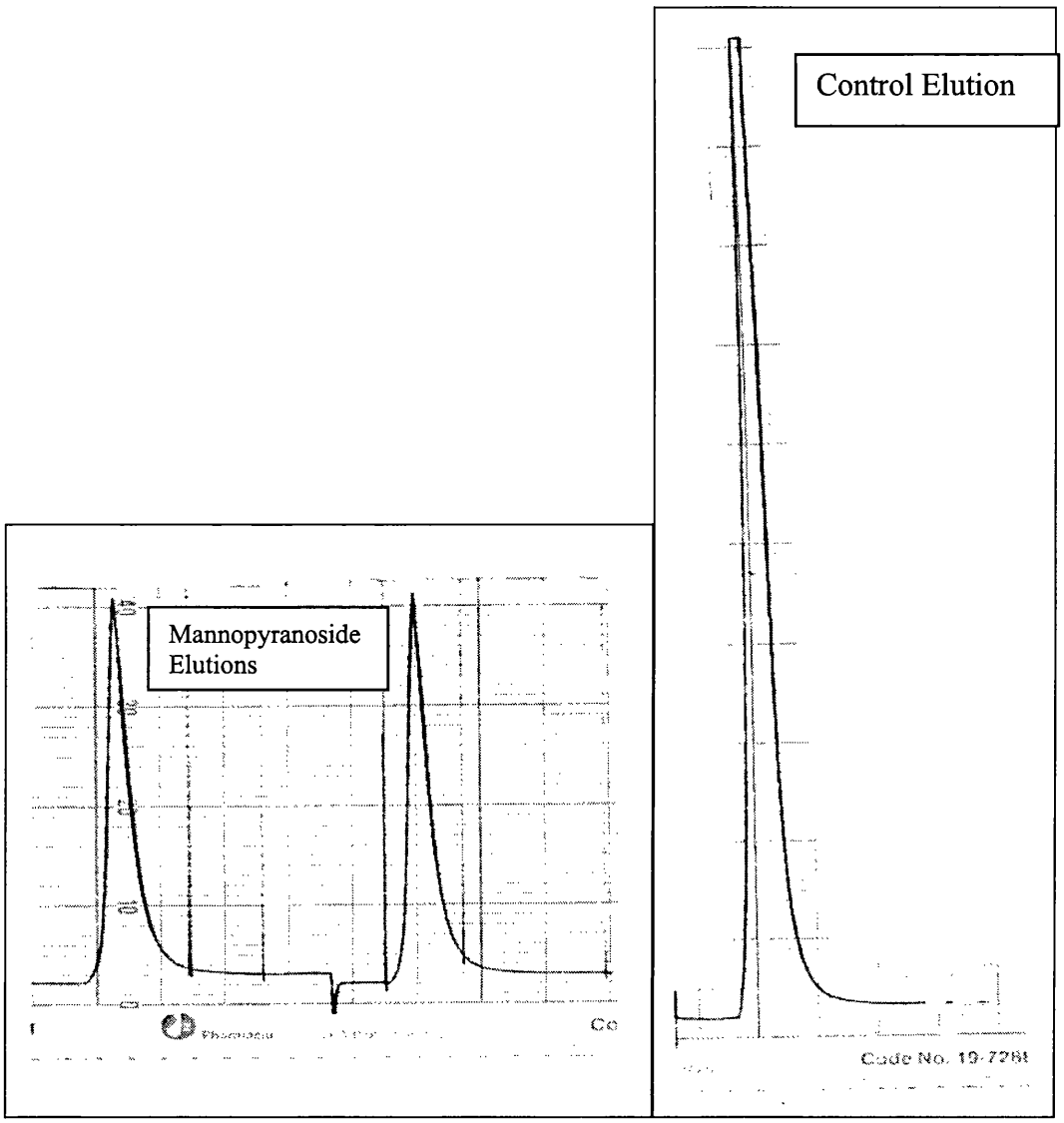


Figure 3.5 Elution profiles using methyl-∝-D-mannopyranoside

Shows the same samples of Figure 3.5 eluted again with methyl-∝-D-mannopyranoside. The obvious peaks indicate the elution of some purified Con A. However when compared to control there is some protein that is still bound to the columns.

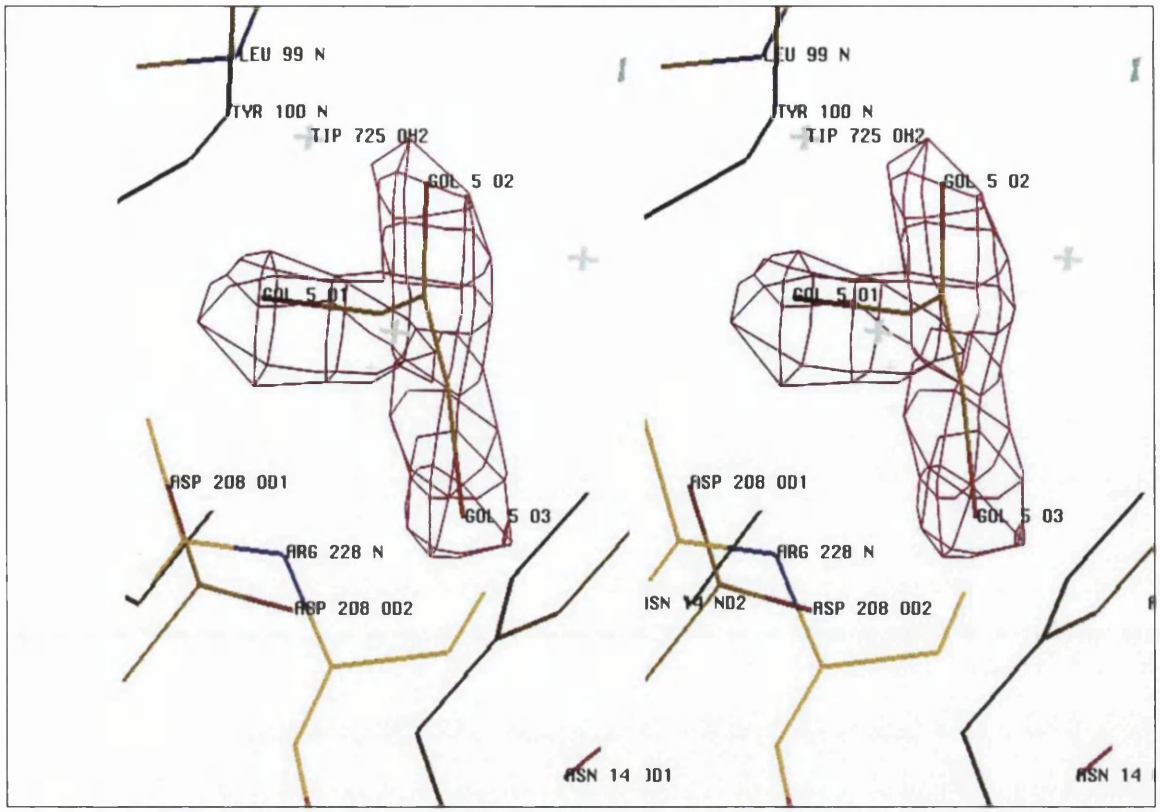


Figure 3.6 Stereo-view of the mFo ÿ Fc map omit map contoured at 5 ¯

Shows the glycerol molecule at G1 was not included in the calculation. Residues involved in carbohydrate recognition and that form hydrogen bonds to the glycerol are labeled. Figure adapted from **(Lo pez-Jaramillo et al., 2004).** Acta Cryst. D60, 1048-1056.

3.1.2 Guanidine Hydrochloride

3.1.2.1 Introduction

Aggregation is a major problem in reducing the refolding yields, thus, the main goals of the current studies were to find a solution to overcome this barrier. Guanidine hydrochloride (Gnd-HCl) as a denaturant in high concentration (8M) is used extensively as part of the isolation and solubilisation process of the insoluble inclusion bodies harvested as bacterial fermentation pellets (refer to section.2.4.5). Some studies indicated the use of this denaturant to reduce the possible formation of aggregates by using it in low concentration, (refer to section 1.4.7.1). Therefore, this current section focused on Gnd-HCL as a refolding aid to improve the overall yield by diluting the primary (8M) Gnd-HCL concentration to (0.27M) which is the refolding concentration currently used in our laboratory. Another trial is also conducted by raising the current Gnd-HCL concentration of (0.27M) during refolding and binding stages to (0.5M). One of the main aims of the current study was to find a suitable Gnd-HCl concentration which can aid the refolding yield to the maximum point and at the same time not affect the binding of the lectin to the dextran affinity matrix.

3.1.2.2 Method

The general method (refer to section 2.4.5) indicated that the initial guanidine hydrochloride concentration present at the start of the refolding process is 0.27 M following the inclusion pellet purification and solubilisation with 8M chaotropic guanidine hydrochloride. The conditions of refolding in the present part were altered by increasing the guanidine-HCl concentrations to 0.5, 0.75, 1.0 and 1.5 M. Figure 3.8 describes the practical scheme used in the design of this section. Furthermore, the effects of guanidine-HCl on binding of active lectin to dextran were also separately assessed so that its usefulness as a folding aid could be determined. The quality of the purified protein was confirmed by sharp elution profiles and spectroscopic wavelength scans. Mean and standard errors of the mean were calculated where applicable.

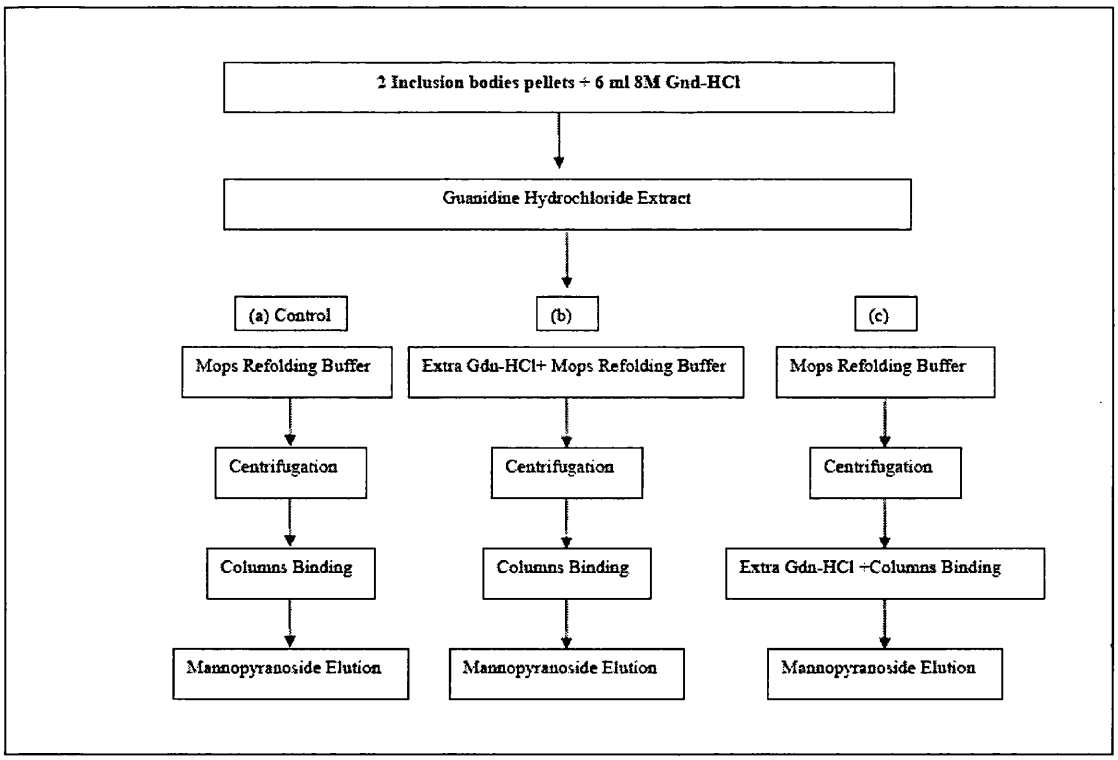


Figure 3.7 Practical steps used in the guanidine hydrochloride study

Shows (a) Control is at 0.27 M Gnd-HCl concentration i.e. without any additions just the guanidine hydrochloride extract used at start to denature the protein. (b) Illustrates the refolding stage with extra-Gnd-HCl, and (c) the binding to the column or the after the refolding stage with extra-Gnd-HCl.

3.1.2.3 Results and Discussion

In this series of experiments (sets of four conditions) guanidine-HCl at concentrations of 0.27, 0.5, 0.75, 1.0 and 1.5 M was investigated during the refolding and binding stages, (Table 3. 2) the first experiment (a) investigated the effect of guanidine hydrochloride at the binding stage. This indicated a small decrease in the purified yield compared to the control when guanidine hydrochloride was used at 0.5 M and this result was confirmed by a similar repeat (b). However, experiment (a) produced a clear decrease in binding when guanidine hydrochloride was used at 1.0 and 1.5 M. Experiment (b) also indicated a clear decrease in yield when guanidine hydrochloride was used at 1.0 and 1.5 M during the refolding stage, (Table 3.2) and (Figure 3.8). Therefore, the effect of guanidine at concentrations less than 1.0 M was investigated in more detail. In addition, two more experiments (c and d) have been conducted with guanidine hydrochloride at 0.5 and 0.75 M during the refolding or binding stages, (Figure 3.19). A distinct increase of 1.23-fold (23 %) in yield compared with the normal procedure (control) is found when 0.5 M guanidine is present during protein refolding. In contrast, there is no noticeable effect (1.0-fold) of 0.5 M guanidine on dextran binding, and a binding decrease (0.78-fold) with 0.75 M. The mean and standard error of the mean indicate the precision of the methods used. The study has found a new thread to follow by adopting the 0.5M Gnd-HCl procedure during refolding in further studies where this concentration lead to greater yields and does not interfere in the binding stage. As a result of the above mentioned it is found that the folding/unfolding process is a complicated not a straight forward issue that proceed through more than one stage. The use of Gnd-HCl at 8M completely unfolds Con A while diluting to 0.27 M allows for refolding. Raising the concentration to 0.5 M permits more refolding to form the native protein, since it can prevent the aggregation to some extent and this result is consistent with the studies reported by (West et al., 1998), about the refolding of carbonic anhydrase and the formation of intermediates and then the native state using 0.23M and 0.6M Gnd-HCL concentration.

Guanidine -HCl (M) added at stage:	Refolding			Binding		
	Yield (mg/l culture)	A_{280}/A_{260}	Relative Yield %	Yield (mg/l culture)	A_{280}/A_{260}	Relative Yield %
0.27 (normal method)	22.2^a,18.7^b 16.3^c, 20.3^d 23.9^c, 23.1^c, 24.8^d M = 21.0 SEM =1.15	1.8^a, 1.9^b 1.8^c 2.4^d 2.0^c, 1.8^c, 1.8^d	100	*22.2^a, 18.7^b, 16.3^c,20.3^d ,23.9^c, 23.1^c, 24.8^d M = 21.0 SEM =1.15*	*1.8^a, 1.9^b 1.8^c 2.4^d 2.0^c, 1.8^c, 1.8^d*	*100*
0.5	20.6^c, 26.3^d 27.9^c, 26.7^c, 27.9^d M = 25.9 SEM=1.36	1.9^c, 2.0^d, 1.9^c, 1.7^c,1.9^d	123.3	20.1^a,17.4^b 17.5^c, 21.1^c, 25.9^c, 24.1^d M =21.0 SEM=1.41	2.3^a, 1.9^b, 1.9^c, 2.2^c,1.9^c4^d	100
0.75	Not Determined	Not Determined	Not Determined	15.0^c,18.3^d M = 16.4	2.2^c 2.0^d	78.1
1.0	10.0^b	1.9^b	47.6	11.7^a	1.8^a	55.7
1.5	12.6^b	1.8^b	60.0	12.6^a	1.8^a	60.0

Table 3.2 Effect of Gnd-HCl concentration on the refolding and binding stages

Control sets were conducted at the 0.27M Gnd-HCl concentration refolding also shown in *italics* for the binding stage. The table indicates an increase of 23.3% in the yield at 0.5 M (refolding) compared with the control during experiments (c and d). Superscript letters (a, b, c, and d) describing the same set of experiments. M= mean, and SEM= standard error of the mean.

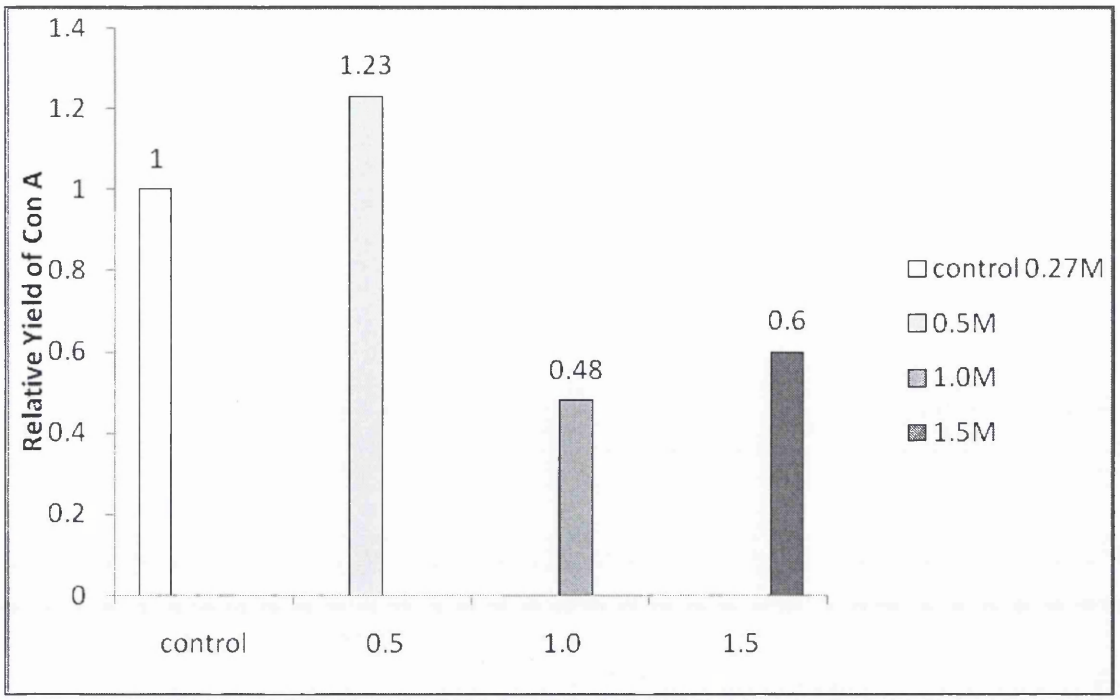


Figure 3.8 Guanidine-HCl Concentration (M) during refolding

Indicates the increase of 23% in yield when Gnd-HCl was added at 0.5 M□ during refolding with significant decrease of 52% and 40% at 1M □ and 1.5 M■ , respectively. Control □ concentration is 0.27 M. All data in Table 3.2 used to plot this Figure.

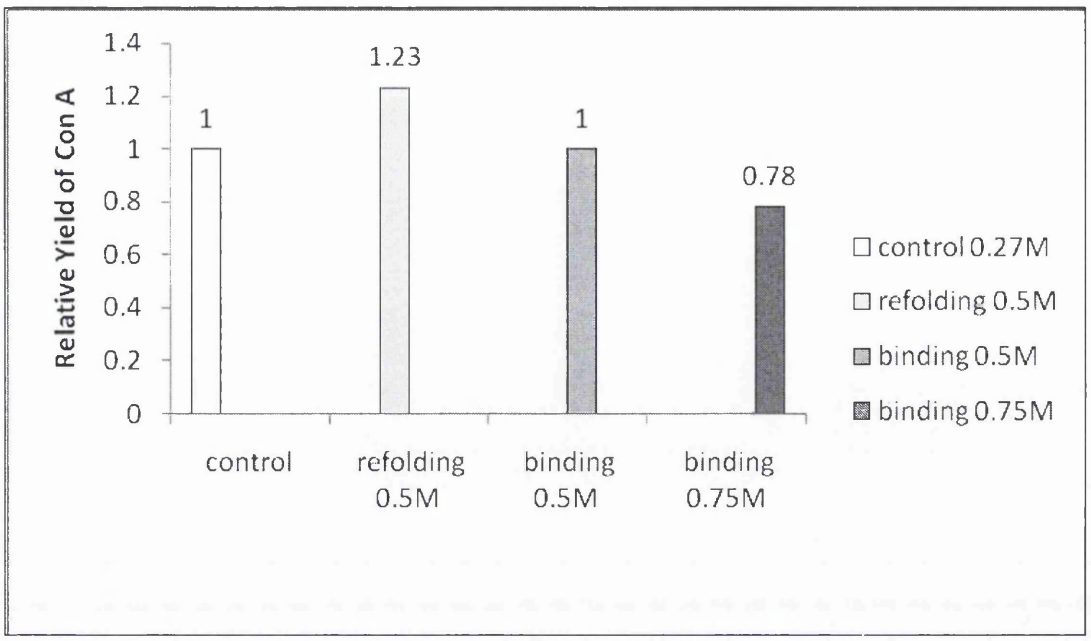


Figure 3.9 Guanidine-HCl Concentration (M) during refolding and binding stages

Shows relative yield of purified Con A (data of experiments c and d only) of Table 3.2 Control ☐ was conducted at 0.27M Gnd-HCL concentration. An increase of 23% is indicated at 0.5 M ☐ refolding, no increase at 0.5M binding ▨ and a decrease of 22% at 0.75 M ▩ binding were observed, respectively.

3.1.3 Dilution

3.1.3.1 Introduction

Dilution was considered as an inexpensive method that can prevent aggregation easily by just pouring the denaturant solution into a suitable refolding buffer (refer to section 1.4.2). The method was extensively used in our laboratory where a previous study had exploited two dilution factors (30- and 60- fold). The aim of this project is to optimise the refolding conditions by finding a suitable refolding environment that assures the production of high renaturation yields. Therefore, picking up the best dilution factor which suits the refolding of Con A is one of these conditions. Dilution was explored by using most of the known protocols like rapid dilution, continuous dilution, pulse dilution and step-wise additions. It is worth mentioning here that mixing is a very important matter in aiding refolding and is greatly associated with dilution (section 1.4.4). Unfortunately, no separate data were available to reveal its value in refolding; however, it has been used extensively in the present refolding studies.

3.1.3.2 Method

The general method stated the practical stages used (refer to section 2.4.5), however, some alterations in the refolding conditions were made in the current sets of experiments. Different rapid dilution factors ranging from 15 fold to 120 fold were investigated using a final Gnd-HCL concentration of 0.5M (conducted by just pouring the denaturant mixture into Mops refolding buffer). Furthermore, other dilution protocols were also explored like continuous dilution where the denaturant solution was added continuously in a drop-wise manner over 1 hour of ice incubation using an automatic pipette to a final Gnd-HCl concentration of either 0.27 or 0.5M. Pulse renaturation protocol was performed by adding portions 6/12= 0.5 ml of the denaturant solution at timed intervals (every 10 minutes) during 2 hours of ice incubation to a final Gnd-HCl concentration of 0.5M. Lastly, the step-wise dilution which is a technique using the rapid dilution method but is carried out by keeping certain time intervals between additions over 1 hour. Two experiments were conducted at different dilution factors 30 fold (3X then after ½ hour 10X) and 15 fold (3X then after ½ hours 5X) aiming towards a final Gnd-HCL concentration of 0.5M. Control sets for the different

experiments were using the rapid dilution method at both Gnd-HCL concentrations (0.27 or 0.5 M). Absorbance ratios of A280/A260 confirm the quality of the purified protein. Means and standard errors were also calculated where applicable.

3.1.3.3 Results and Discussion

The aim of the current series of studies as was mentioned above was to tackle aggregation by either preventing or reducing its formation. Most of the documented literature has focused on the elimination of aggregation by reducing the protein concentration depending on the existence of a strong relationship between protein concentration and the presence of aggregates, thus, insisting that the prevention of the competition between aggregations and refolding is a vital matter. Therefore, here in this section the various dilution protocols like rapid, continuous, pulse, and step-wise additions using two guanidine-HCl concentrations of 0.27 and 0.5 M (where applicable) were deeply investigated. Firstly, a study illustrating the effect of different rapid dilution factors on the refolding of Con A was conducted. Table 3.3 and Figure 3.10 compare the relative yields at different dilution factors (15X -120X). As a result, dilution factor 30 fold at 0.5M concentration appeared as the optimum factor, since, an increase of 14% to the control has been recorded and the quality of the protein represented by the A280/A260 ratios was good. Furthermore, (Figure 3.11) indicates a relationship between the protein concentration and the renaturation yields. If it is assumed that the control refolds all the protein present so that 27.3 mg/l culture corresponds to the total concentration of Con A, this may be converted to concentration in μg/ml that would be present in the 180 ml in the refolding flask at 30X dilution. Then the concentration of Con A at 15X is twice this, concentration at 60X is half this, etc. The bell-shaped curve shows the type of relationship found.

Secondly, another study investigated the role played by different dilution protocols on the prevention of aggregate formation (Table 3.4 and Figure 3.12) compared between controls (rapid dilution), continuous, pulse, and step- wise additions. Continuous renaturation at (0.5M) Gnd-HCl concentration has shown an increase of 10% during 1 hour of ice incubation whereas the pulse renaturation method indicated a decrease of 23% when incubated for 2 hours. On the other hand, there is a negligible decrease of 3% when continuous renaturation was compared to control at (0.27M) during1 hour of ice

incubation. Figures (3.13 and 3.14) show the gradual increase in the Gnd-HCl concentration during the two protocols used (the continuous dilution and the pulse renaturation method). The striking feature in this study was the very low yield obtained using the step-wise dilution (Table 3.4) that clearly indicated the presence of aggregates by the obvious turbidity (a visual observation) of the refolding mixtures when compared to the controls. This result implied that the starting high concentration of both the denaturant and the protein (Table 3.5) with an initial protein and denaturant concentration of 620μg/ml and 2.9 M respectively was the main cause of aggregation. Thus, the step-wise method is not one of the favoured ways for the refolding of Con A. (Terashima et al., 1996) reported that the rate by which the denaturant should be added to the refolding buffer must be slower than the rate determining the folding step for the protein to prevent aggregation. The rapid dilution protocol results indicate the crucial role played by the Gnd-HCl concentration and the volume of the refolding buffer solution when the start point was at high concentration of both Gnd-HCl and the protein. The occurrence of sudden low Gnd-HCl concentration during the first part of the refolding process enable formation of the first partly folded forms and the diluted protein concentration favours the prevention of aggregates. (Figure 3.15) indicates a hypothetical pathway describing the events of the formation of a native soluble active tetramer of Con A when starting at high Gnd-HCl and protein concentrations since aggregation at high protein concentrations leads to low yields. However, low protein concentration also reduces the yields as seen in Figures (3.10 and 3.11). (Figure 3.15) suggests also that the formation of an unstable monomer is the determining factor for successful refolding as it cannot fold independently before associating but needs a partner at an early stage of the refolding process to form a dimer. (Figure 3.15) can represent all sorts of other dilution protocols in addition to the rapid dilution method. These ideas are also illustrated in (Figure 3.16).

To conclude with, high protein concentration reduces the overall refolding yields due to the formation of aggregates. However, very low protein concentrations can lead also to other problems where the possible difficulty for the monomer to find a partner also reduces yields. The presence of a stable monomer would lead to fully folded dimer as it is less sensitive to dilution and with unlimited time to find a partner. Tetramer is then formed at the end of the process by the reversible association of fully folded dimers. On the contrary, the unstable monomer is very sensitive to its environment and without finding a partner in a limited time it may become part of an aggregate or otherwise

misfold. This misfolded monomer might be another inactive but soluble form of monomer that is unable due to its altered structure to bind to a partner and form a partially folded dimer. The critical element of this scheme is the suggestion that Con A is more likely to fold via the formation of dimers early in the process so that unstable monomers are recruited into more stable dimers rather than being able to fold in isolation as stable monomers. If there is a limited time for an unstable subunit to find a similar partner before it becomes unable to bind it, then low protein concentrations will discourage correct folding. In other words, it is proposed that early dimer formation is essential for folding to proceed correctly, rather than that late dimerisation occurs after the correct folding of isolated monomers is complete.

The accurate determination of the appearance of the hypothetical partially folded unstable monomer might open a new door for refolding in industry. This work might help overcome the barriers when refolding by dilution through suggesting economical ways to control the refolding buffer volumes and the equipment needed for the process.

DILUTION FACTOR (X)	Yield (mg/l culture)	A_{280}/A_{260}	Relative Yield %
15	14.3^a 17.3^b 18.8^c M = 16.8 SEM = 1.32	2.3^a 1.9^b 2.0^c	61.5
20	22.8^d	2.2^d	83.5
30	28.0^a 26.7^b 27.9^c 26.6^d M = 27.3 SEM = 0.45	1.9^a 1.8^b 1.8^c 2.3^d	100
60	25.5^c, 24.6^d M = 25.1	1.7^c 2.3^d	91.9
120	19.1^d	1.3^d	70.0

Table 3.3 The effect of rapid dilution factors on the refolding stage at 0.5 M Guanidine HCl

Table shows the effect of different dilution factors at 0.5 M Gnd-HCl on the overall refolding yield. Dilution factor 30X giving the optimum yield mean=27.28 ± 0.45. Superscript letters a, b, c, and d indicates a specific experiment with the purity of examined sample shown by corresponding absorbance ratio.

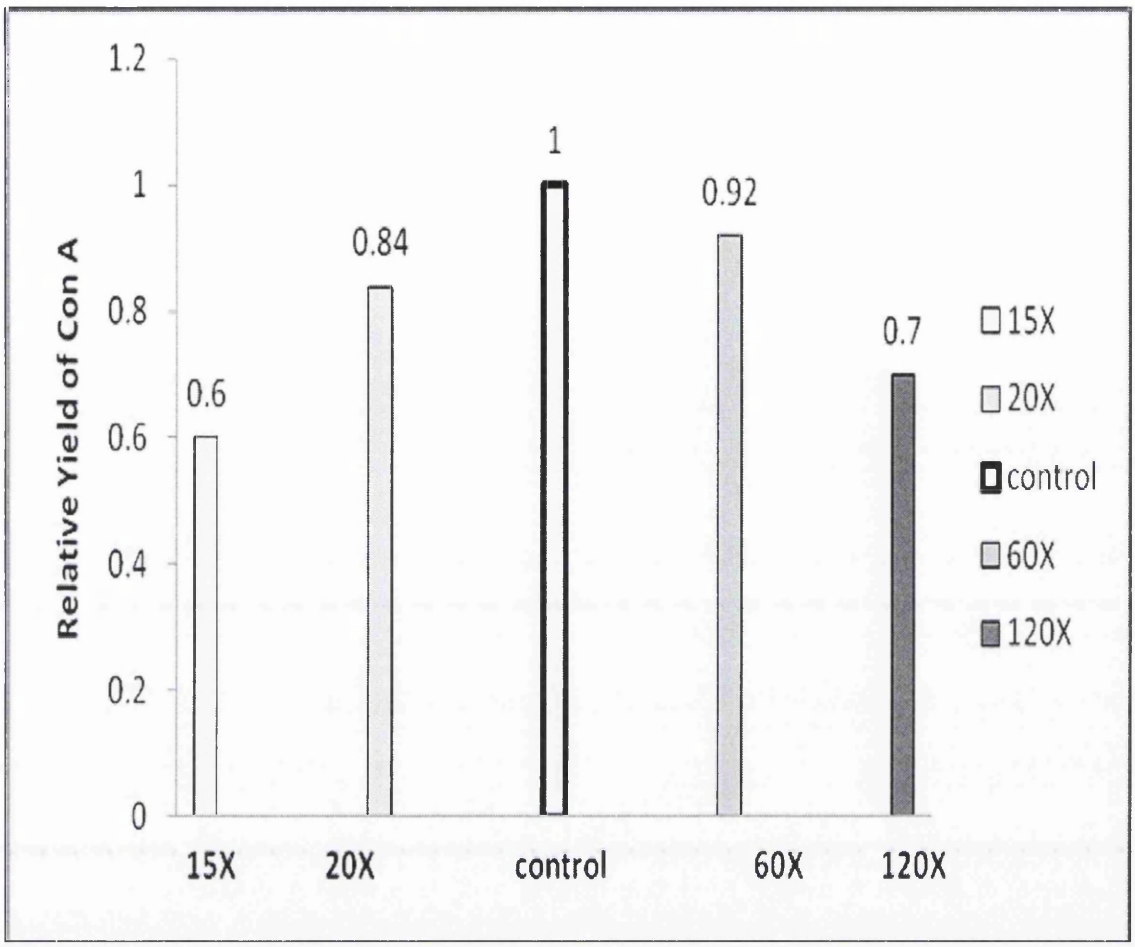


Figure 3.10 Effect of different (rapid) dilution factors on the refolding stage of Con A

Describes the effect of different dilution factors on the refolding stage of Con A at 0.5M Gnd-HCl concentration. Histogram shows an approximate bell shape where, Control sets □ at dilution factor 30X show maximum yield, and sets of dilution factor □ 15X indicate the minimum. Figure used data from Table 3.3.

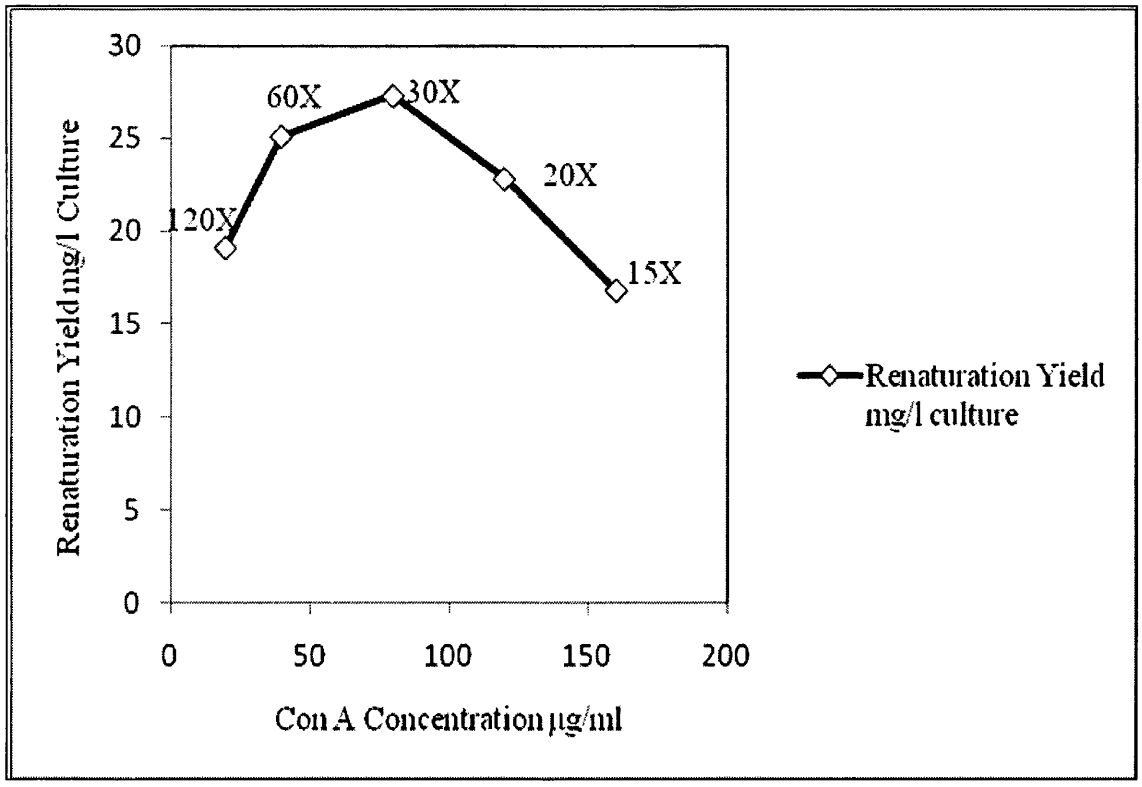


Figure 3.11 Effect of protein concentration on the renaturation yield (At 0.5M Gnd-HCl concentration using rapid dilution)

Relates the protein concentration to the overall renaturation yield when refolding at 0.5M Gnd-HCl final concentration and using the rapid dilution method. The bell shape curve indicates the optimum protein concentration as 80 µg/ml according to 27.3 mg/l culture yield. Protein concentrations ≤50 µg/ml provides lower renaturation yields.

Dilution Protocol										
Refolding over 3 - 4 hours:										
Gnd-HCl concentration	Rapid Dilution (30X) (Control) 1 hour on ice, 2 hours on bench		Continuous dilution (≤ 30X) During 1 hour on ice, 2 hours on bench		Pulse dilution (≤ 30X) During 2 hours on ice, 2 hours on bench		Step dilution (3×, 30 min then 10×) 1 hour on ice, 2 hours on bench		Step dilution (3×, 30 min then 5×) 1 hour on ice, 2 hours on bench	
	Yield (mg/l culture)	A280/ A260	Yield (mg/l culture)	A280/A260	Yield (mg/l culture)	A280/ A260	Yield (mg/l culture)	A280/ A260	Yield (mg/l culture)	A280/ A260
(0.5M) Gnd-HCL	24.5 21.5 21.6 M=22.5 SEM= 0.98	2.2 1.9 3.0	21.6 26.2 24.4 22.0 M=23.6 SEM= 1.08	3.0 1.9 1.8 1.7	17.5	2.1	2.0	4.1	0.5	3.2
(0.27M) Gnd-HCL	22.2	1.8	22.6 21.0 M=21.8	1.7 2.9	Not Determined	–	Not Determined	–	Not Determined	–

Table 3.4 Effect of different dilution protocols on the refolding stage (At (0.27 and 0.5) M guanidine –HCl concentrations)

Results of the different dilution protocols using 0.27 and 0.5 M Gnd-HCl final concentrations over 1 - 2 hours ice incubation. The table compares controls, continuous, pulse and / or step- wise dilution, where continuous renaturation (0.5M) is showing an increase of 10% after ice incubation for 1 hour with a pulse dilution decrease of 23% when incubated for 2 hours. On the other hand, there is a negligible decrease of 3% when continuous renaturation was compared to control at 0.27M with 1 hour ice incubation. Step dilution in dilutions (3×, 30 min then 10×) or (3×, 30 min then 5×) shown the least yields of all, where much greater turbidity of buffer solutions was the dominant feature (visual observation). Mean (M) and Standard Error of the Mean (SEM) indicated where applicable.

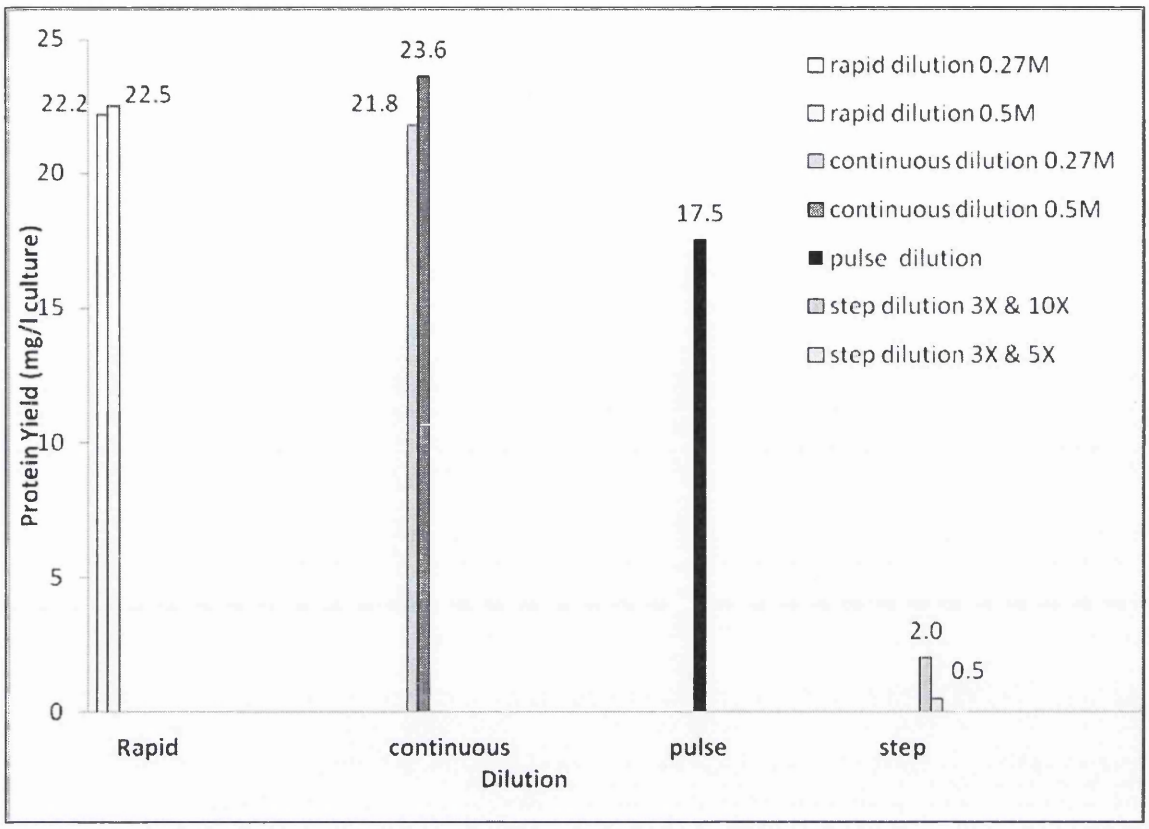


Figure 3.12 Effect of various dilution protocols on the refolding stage of Con A

Illustrates the effect of different dilution protocols on the refolding stage at two Gnd-HCl concentrations (0.27 and 0.5) M. control sets are rapid dilution at factor (30X). Continuous dilution ▩ during 1 hour and at 0.5M Gnd-HCl level indicates the highest refolding yield. Figure used data from Table 3.4.

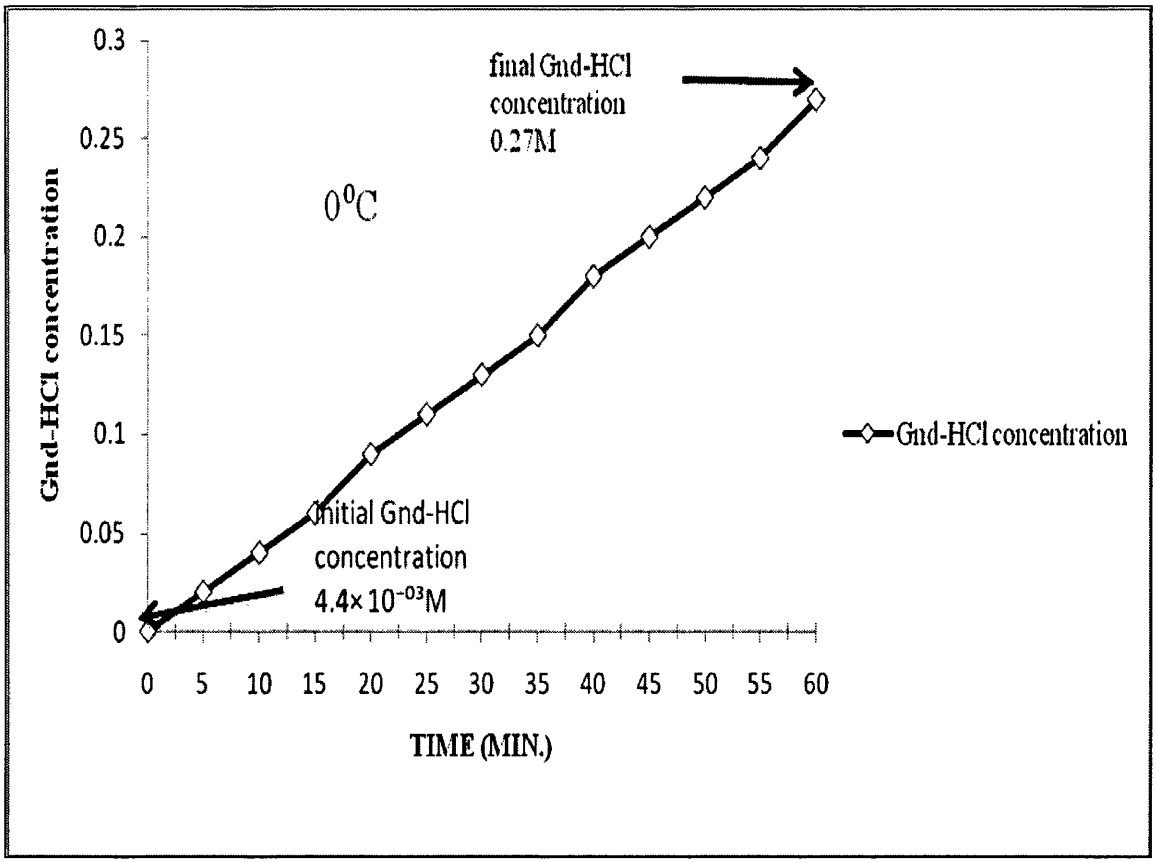


Figure 3.13 Gnd-HCl concentration during 1 hour ice incubation period of the refolding process (Using continuous dilution protocol)

Shows the gradual increase in Gnd-HCl concentration from an initial concentration of 4.4×10^{-3} to a final concentration of 0.27M. Additions were determined as drop sizes of ~0.1ml/min. and refolding buffer volume =180 mls. Based on Data taken from Table 3.4.

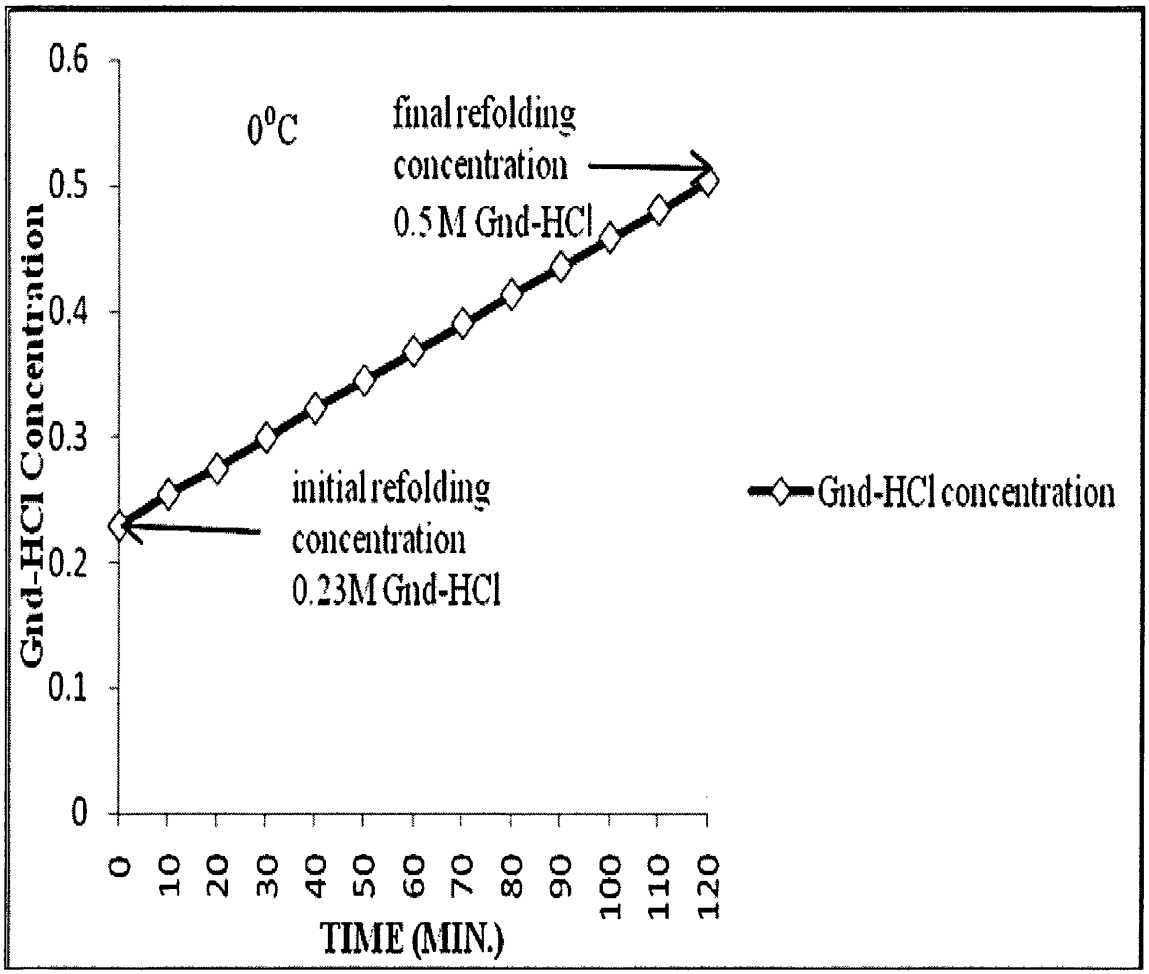


Figure 3.14 Gnd-HCl concentration during 2 hrs ice incubation period of the refolding process (using pulse dilution protocol)

Shows the gradual increase of Gnd-HCl concentration to final concentration of 0.5M during the 2 hours ice incubation indicating an initial refolding buffer concentration of Gnd-HCl = 0.23M. Additions of Denaturant solution are timed by every 10 min. with drop size = 0.5ml/10min. Hence, denaturant concentration of~0.023M is added every 10 min. Based on Data taken from Table 3.4.

Protocol	Control (Rapid Dilution)		Step dilution (3×,30min.then 10×)		Step dilution (3×,30min.then 5×)	
	Initial Concentration	Final Concentration	Initial Concentration	Final Concentration	Initial Concentration	Final Concentration
Protein Concentration (μg/ml)	62	62	620	62	620	124
Denaturant Concentration (M)	0.5	0.5	2.9	0.5	2.9	1.0

Table 3.5 Denaturant and protein concentrations during 1 hour ice incubation stage of the step dilution

Indicates the Initial and final protein and denaturant concentrations. Control at 0.5M Gnd-HCl concentration and conducted using the rapid dilution method (1 hour on ice, 2 hours on bench). Step Dilution 30X and Step Dilution 15X started both with initial protein and denaturant concentrations of 620μg/ml and 2.9M respectively. Based on Data taken from Table 3.4.

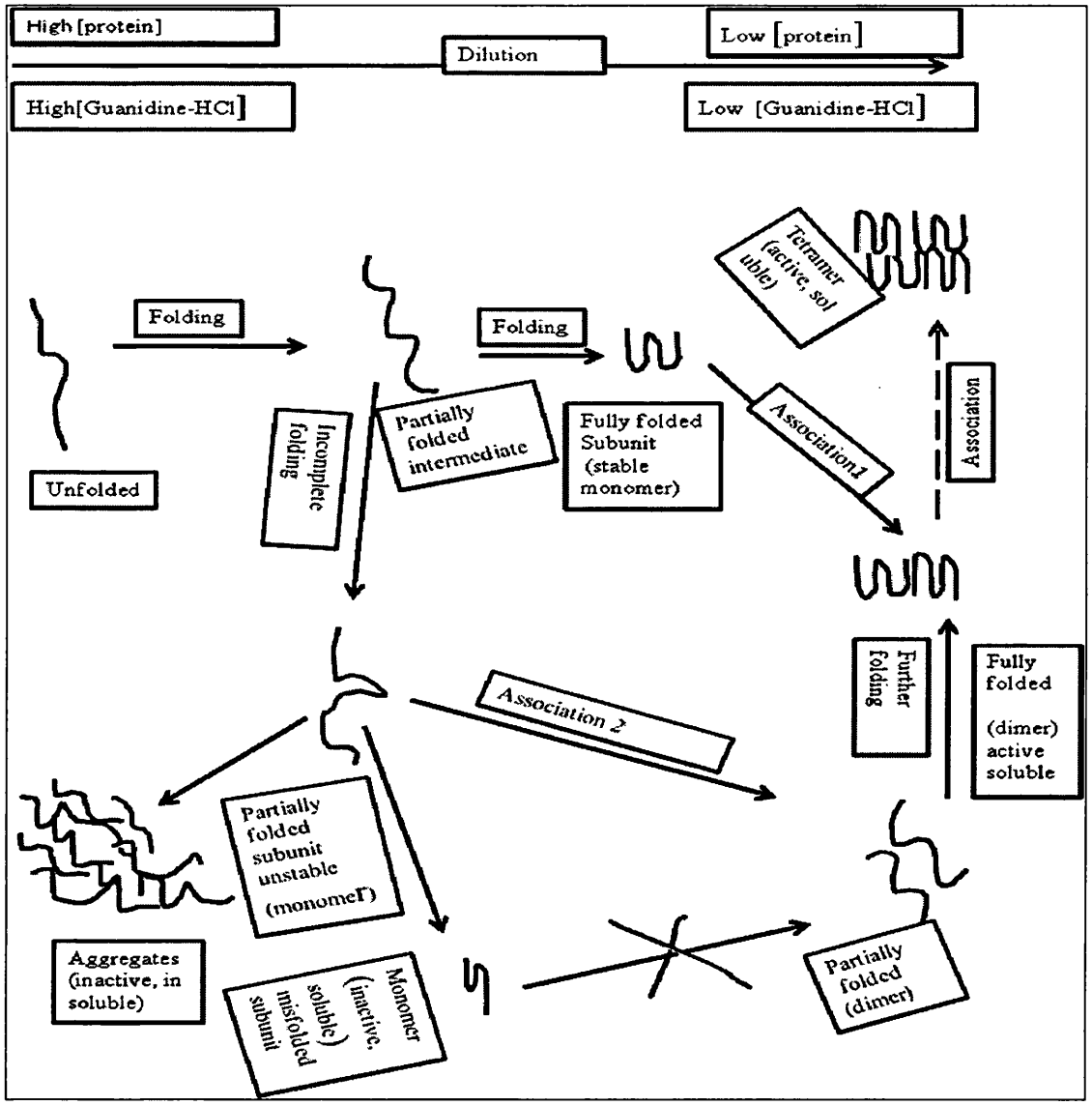


Figure 3.15 Hypothetical pathway of the effect of dilution on the formation of native Con A tetramer

Shows the effect of varying concentrations of Gnd-HCl and protein on the formation of biologically active Con A and also aggregates.
The presence of the partially folded unstable monomer in favourable refolding conditions is assumed to determine the success of the renaturation process.
1-Stable monomer association is less sensitive to dilution since unlimited time to find partner.
2-Unstable monomer association is more sensitive to dilution since limited time to find partner to continue folding successfully.

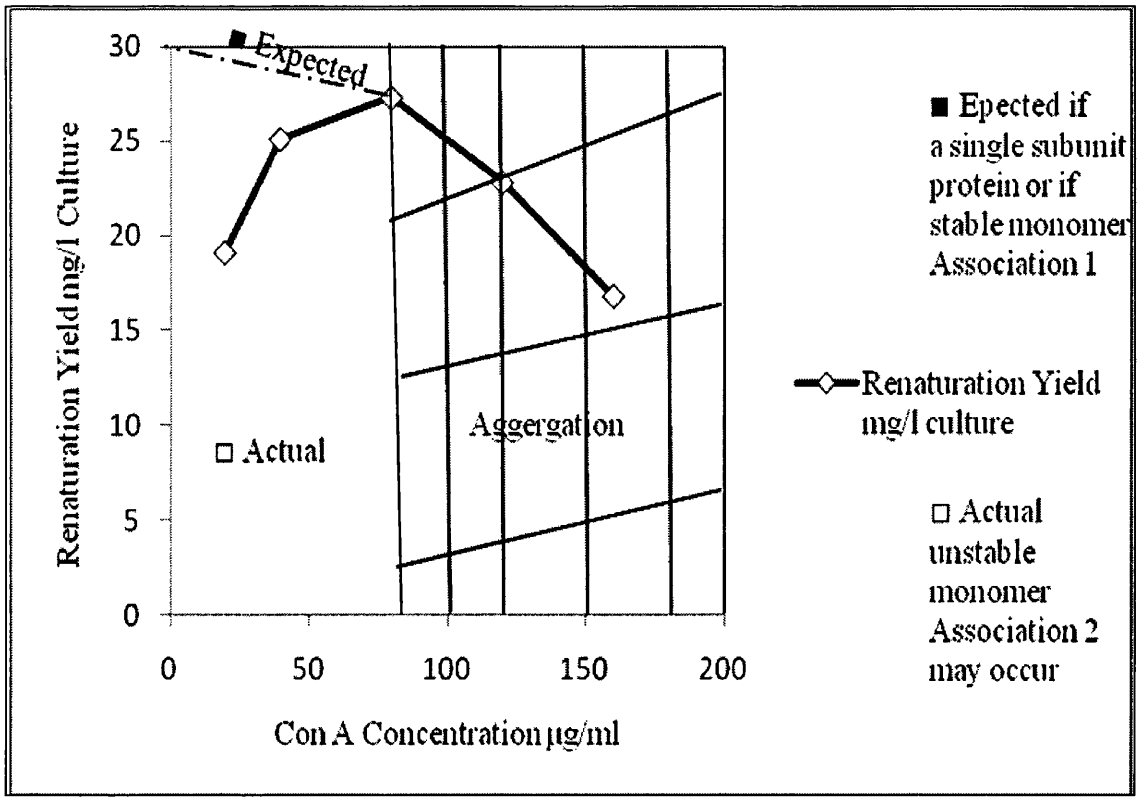


Figure 3.16 Hypothetical events during the refolding of Con A

Shows three hypothetical events during the refolding of Con A relating to Figure 3.16, where Aggregation here comprises the right-hand side: partially folded unstable monomer leads to aggregates at high protein concentrations. Expected represents the association 1 pathway where fully folded monomer as a stable intermediate gives rise to fully folded dimer. Actual represents Association 2 pathway where the formation of a partially folded dimer by an unstable monomer is more sensitive to lowered protein concentration.

3.1.5 Temperature

3.1.5.1 Introduction

The present study concentrated on the role of incubation temperature during the refolding stage as a potential factor in optimising the refolding conditions. As the goal for the whole project is the production of improved yields by preventing the aggregation process, temperature was found to play a leading role in this aspect (Section 1.4.6) The current protocol of methodology used in the laboratory (section 2.4.5) always maintain the refolding solution at 0 °C (ice incubation) then a warming up step by allowing standing on bench temperature incubation (this, however, depends on the season of the year whether it is warm or cool). This part will try to explore the various temperatures and their effects on the refolding and binding stages and the role which is played by them in reducing the aggregation process. Furthermore, observations are made at two different Gnd-HCl concentrations combined with the incubation temperatures on the refolding process.

3.1.5.2 Methods

The general method (refer to sec.2.4.5) indicates that the initial guanidine hydrochloride concentration present during the refolding process is 0.27 M. The present study investigated the role of the refolding incubation temperature by changing the design of the general refolding experiments as follows: (a) refolding over 3 hours using constant temperatures of -5.5, 0, 4, 15, 25, or 37°C through out the refolding process. (b) Refolding over 3 hours: that is 2 hours ice incubation and 1 hour at 0, 4, 15, or 25°C. The control refolding conditions for both (a) and (b) were; 2 hours on ice and 1 hour standing on the bench. Temperatures were maintained as follows: 15, 25, or 37 °C by using constant temperature water baths, 0°C by using an ice bath, 4°C by using the constant temperature cold room, and finally -5.5°C by using an ice-salt bath (a mixture of ice and 5% w/w sodium chloride) as the salt causes the mixture to absorb more energy from the environment (becoming colder). The -5.5°C condition was considered the lowest practicable temperature to reach before the freezing point of the Gnd-HCl refolding buffer (-7.4°C). The Gnd-HCL concentration in this study is at 0.5M and the quality of protein purified is measured with the absorbance ratios of (A280/A260)

3.1.5.3 Results and Discussion

The current study aimed at the prevention of aggregation through the control of the incubation temperature during the refolding stage. Several temperatures were selected for trial so as to optimise the refolding conditions. (Table 3.6) indicates the design of two experimental sets, the first was conducted by using constant selected temperatures of -5.5, 0, 4, 15, 25, or 37°C over 3 hours of the refolding stage, the second was conducted in the normal way i.e., 2 hours ice incubation and then 1 hour incubation at one of the previously mentioned temperatures. Part (a) of (Table 3.6) reports decreases of 15.9, 54.80, 45.34, 100, and 100 % at -5.5, 4, 15, 25, and 37°C, respectively, compared to control sets. Part (b) of (Table 3.6) also indicates decreases of 21.31, 18.90, and 59.75 % at 4, 15, and 25°C compared to the control sets. However, comparisons between (a) and (b) indicates increases of 74.12, 48.38, and almost 100 % at 4, 15, and 25°C, respectively. This is similar to the findings of Wetlaufer and Xie (1996) that their temperature-leap technique (Figure 3.19) can improve yields compared to constant temperature incubation throughout. These workers also saw decreasing yields as temperatures were increased in their experiments (Figure 3.20). Refolding with continuous incubation at 0°C produced the maximum yield as shown in Figures (3.17 and 3.18). However, subsequent re-appearance of turbidity of these 0°C refolded mixtures during affinity column loading was noticed and this indicated the presence of some further aggregation problems. Aggregation and precipitation during overnight affinity column loading threatens column blockage that can initiate leaking of chromatography columns and loss of refolded protein. Although, Yoshii et al. (2000) dismissed the use of low temperature as it slowed down the refolding process in their studies, this was not the case here as good yields were recovered at low temperatures. The present study agrees with them in that low temperatures have a definite role in the prevention of hydrophobic aggregation. So, for practical uses, a rise in temperature was found to be important for the prevention of precipitate formation during column loading and to eliminate this problem after refolding at 0°C. Therefore, a bench warm up step was adopted (allows for gradual warm up of refolding solutions from 0 to 10-13°C during 1 hour as measured with an electronic thermometer). The present study also used water baths to reach certain temperatures more quickly as shown in (Figure 3.18).

However, the "on bench" incubation protocol allowed for the recovery of better renaturation yields because of the gradual increase in temperature when the flask is not immersed in a water bath. Refolding during 3 hours incubation at 25 and 37°C produced zero yields, high turbidity and large accumulations of precipitate showing the unfavourable effects of high temperatures on refolding. A low temperature at the early stages of folding may improve the monomer's stability allowing more time for association into the dimer structure of Con A, as well as decreasing the hydrophobic interactions responsible for aggregate formation.

(a) Refolding conditions 3 hour continuous incubation at given temperature	Yield (mg/l culture)	(A_{280}/ A_{260})	Relative Yield %	(b) Refolding conditions 2 hours incubation on ice, then 1 h at given temperature	Yield (mg/l culture)	(A_{280}/A_{260})	Relative Yield %
Control 2 hours on ice, 1 hour on bench	26.0, 27.6, 28.0, 25.1 M= 26.7 SEM=0.71	1.8, 2.8, 2..2, 2.3	100	*Control 2 hours on ice, 1 hour on bench*	*26.0, 27.6, 28.0, 25.1 M=26.7 SEM=0.71*	*1.8,2.8, 2..2, 2.3*	*100*
-5.5°C	22.6, 22.2 M= 22.4	2.1, 2.1	83.9	-5.5°C	Not Determined	Not Determined	Not Determined
0°C	24.5 29.0 M =26.8	2.0 2.5	100.4	0°C	*24.5 29.0 M = 26.8*	*2.0 2.5*	100.4
4°C	12.1	3.0	45.3	4°C	19.6, 22.4 M= 21.0	2.7, 2.5	78.7
15°C	14.8, 14.5, 14.5 M= 14.6 SEM=0.02	1.8, 1.9, 2.1	54.5	15°C	23.0, 20.2 M= 21.6	1.9, 2.3	80.9
25°C	0.0	-	0.0	25°C	12,10.4, 9.8 M= 10.7 SEM=0.75	2.4, 2.0, 2.5	40.1
37°C	0.0	-	0.0	37°C	Not Determined	Not Determined	Not Determined

Table 3.6 Refolding conditions at different incubation temperature sets

Shows the refolding conditions over 3 hours where (a) illustrates an incubation period of 3 hours at given temperature and (b) a 2 hours incubation period on ice then 1 hour at given temperature. Control sets were conducted as 2 hours on ice incubation, and then 1 hour on bench temperature. No detectable yields were obtained when refolding at constant temperatures of 25°C and 37°C, while 0°C produced the maximum yield. Refolding conditions of 2 hours on ice incubation then 1 hour at the given temperature indicated 0°C as having the maximum yield and enhanced by 40.1% of the relative yield at 25°C compared to the control.

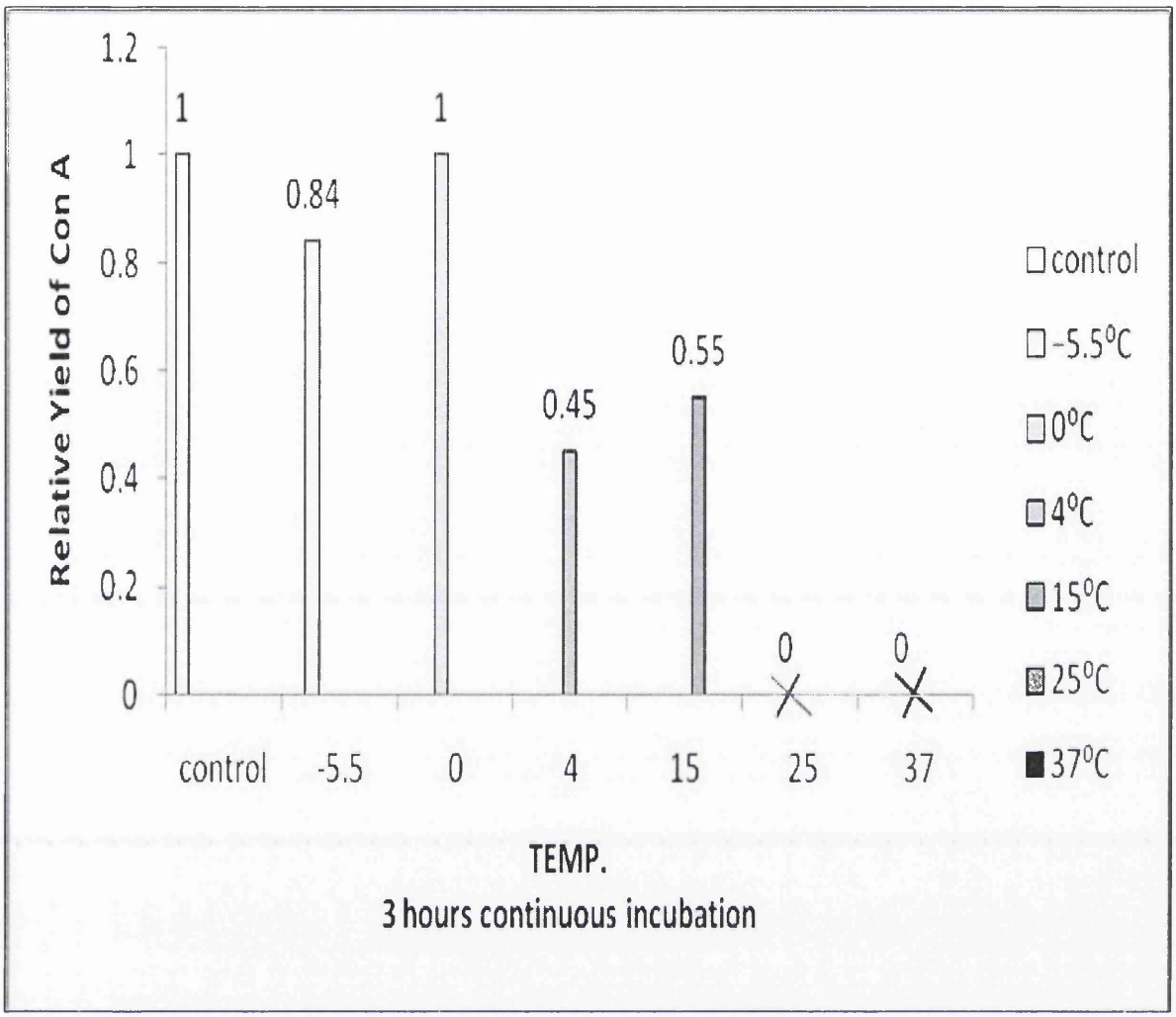


Figure 3.17 Relative yields of Con A after 3 hours of continuous incubation (At different temperatures)

Shows the relationship between the relative yields after 3 hours incubation at different constant temperatures. Controls were 2 hours ice incubation and 1 hour on bench incubation. 0°C gave approximately the same relative yield as the control while other temperatures of -5.5, 4, and 15 °C had decreased relative yields of 0.16, 0.55, and 0.45, respectively. Temperatures of 25°C and 37°C gave zero yields. Figure used data from Table 3.6, part (a).

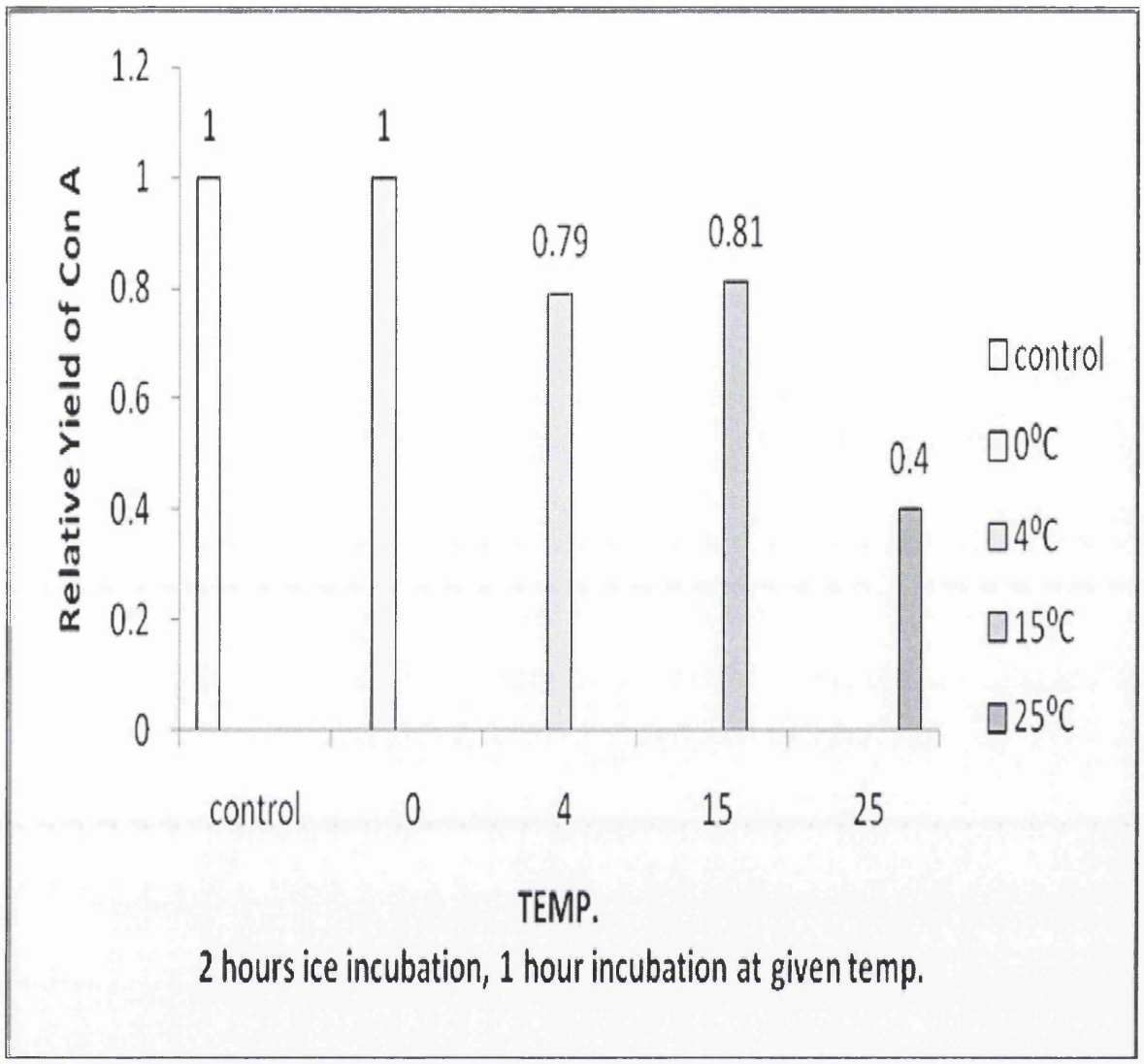


Figure 3.18 Relative Yields of Con A after 2 hours of Ice Incubation and 1 hour (At given temperatures)

Shows the refolding incubation periods of 2 hours on ice then 1 hour incubation at different given temperatures. Figure indicates the same relative yield for the 0°C and control, also between 4 and 15 °C with a decrease of 20 % and 25°C with a decrease of 60% compared to the control. Figure used data from Table 3.6, part (b).

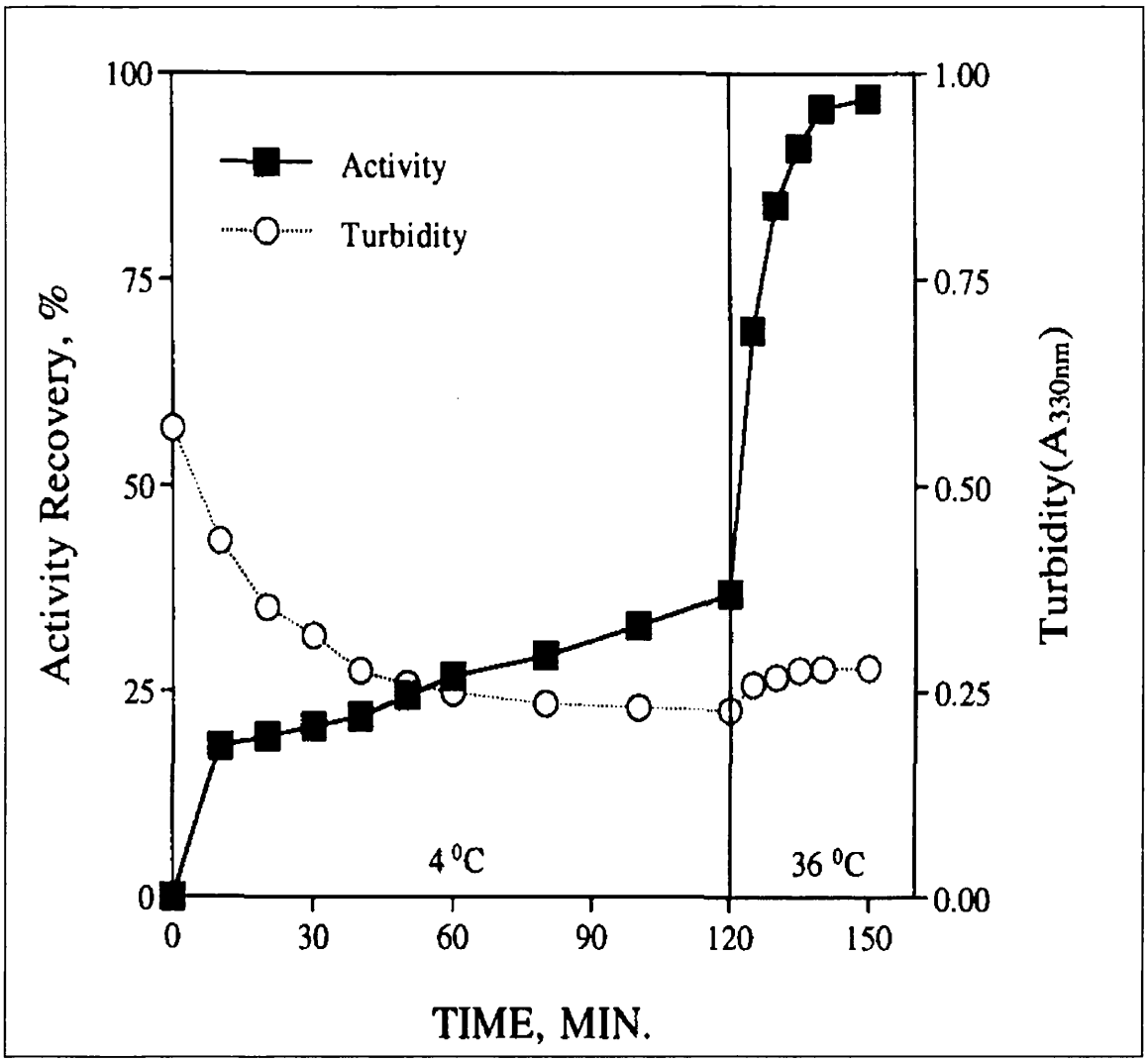


Figure 3.19 Temperature-leap refolding of bovine carbonic anhydrase II (At high protein concentration (4.0 mg/mL), 1.00 M GdmCl)

The first 120 min of the refolding was carried out at 4°C; the sample was then transferred to a 36°C thermostated bath. Activity assays and turbidity measurements were made on separate portions of the same sample. Figure adapted from **XIE Y., and WETLAUFER D. B., (*1996*).** Protein *Sci.* **5**: 517-523.

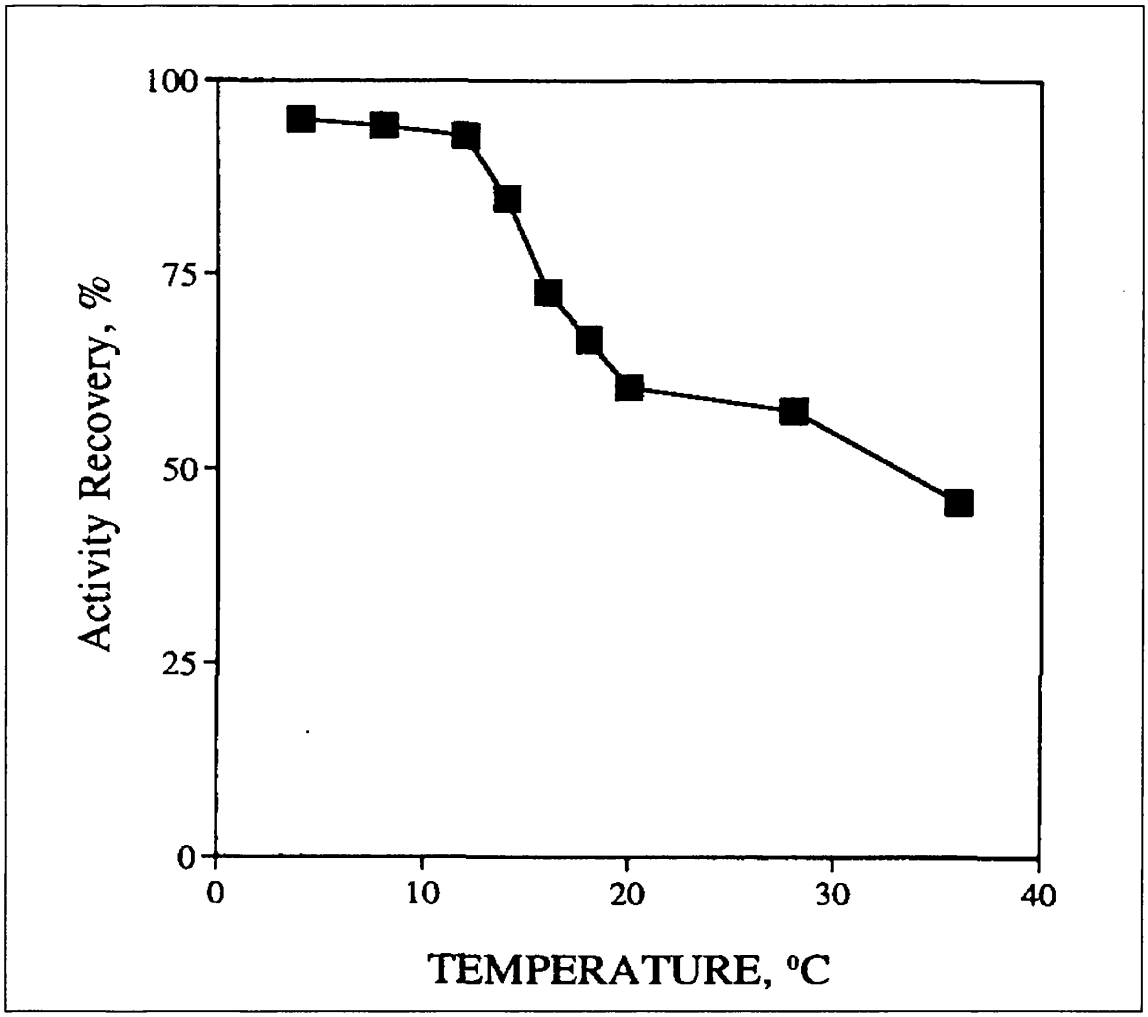


Figure 3.20 Activity recovery at 150 min in a series of temperature-leap experiments (on bovine carbonic anhydrase II)

(As in the previous Figure), where the first 120 min refolding is conducted at the temperature plotted; the sample is then transferred to a 36°C thermostated bath and the activity assayed at 150 min. Figure adapted from **XIE Y., and WETLAUFER D. B, (*1996*).** Protein *Sci.* **5**: 517-523.

3.1.4 TRITON X-100

3.1.4.1 Introduction

The first step for a successful protein renaturation is to prevent aggregation through de-tangling of the mis-folded or aggregated particles. Detergents (section 1.4.7.3) were formerly used in protein biochemistry mainly for the solubilisation and extraction of membrane proteins during purification. However, a report by Zardeneta and Horowitz (1994) has reviewed the efficacy of non-ionic detergents like Tween 20, Triton X-100 and Octylglucoside in assisting the refolding of rhodanese. Triton X-100 is a mild non-ionic detergent and is the member of the Triton family that is extensively used for its efficiency in membrane solubilisation and is less costly than any other detergent from the same family (Slinde and Flatmark, 1976). In the present series of studies, Triton X-100 as a possible refolding assistant is investigated for its expected role in preventing or inhibiting the aggregation process with consequent yield improvements. Aggregation at high protein concentrations was shown to be a problem in the previous dilution experiments section. Wetlaufer and Xie (1996) determined a suitable concentration of Triton X-100 (ca.10 mM) that completely inhibited the aggregation and improved the refolding yield of carbonic anhydrase. The plan of the present study allowed for the exploration of all the previously mentioned aspects.

3.1.4.2 Methods

The general method (refer to section 2.4.5) indicated that the initial guanidine hydrochloride concentration is 0.27 M at the start of the refolding period and also the periods of ice incubation and the standing on bench incubation. In the following part, two experimental designs were planned using Triton X-100 at a concentration of 0.5 % (v/v) as a potential refolding aid (Figure 2.2).

The first was to investigate the role of Triton X-100 on the refolding and binding stages using 0.5 M Gnd-HCl concentration. This experiment was planned as: a) the effect of TritonX-100 added on the start of the refolding process, b) after 2 hours of ice incubation, and c) before the binding to chromatography columns (i.e. after 2 hours of ice incubation, 1 hour standing on bench incubation, and then spinning).

The second study investigated the way in which Con A refolds and the time relationship associated with the refolding process. Therefore, the experimental design was additions of Triton X-100 (a) using 0.5 M Gnd-HCl concentration at: 0 hour, 1/2 hour on ice, 1 hour on ice, 2 hours on ice, 4 hours (2 hours on ice then 2 hours standing on bench incubation). (b) Using 0.27 M Gnd-HCl concentration (for comparison reasons between Gnd-HCl concentrations), Triton X-100 was added at: 0 hour, 1 hour on ice, and 2 hours (1 hour on ice then 1 hour standing on bench temperature incubation). In addition, an extreme experiment using 0.5M Gnd-HCl concentration (all night on ice incubation and all day on bench incubation) was used to explore further the time-dependence of the refolding. Control sets are always maintained at the general refolding conditions with no addition of Triton X-100 and in this study designed using either 0.27 M or 0.5 M Gnd-HCl concentrations. The quality of the protein produced is represented by high A_{280}/A_{260} ratios.

3.1.4.3 Results and Discussion

In this series of experiments, Triton X-100 at a concentration of 0.5% (v/v) was investigated during the refolding and binding stages at a final Gnd-HCl concentration of 0.5 M (Table 3. 7). The first experiment (a) Triton X-100 addition at the start of the refolding period indicated zero yields compared to the control during the refolding stage. The second experiment (b) Triton X-100 addition after 2 hours ice incubation indicated a decrease of 51.61% in the yield compared to the controls. The third experiment (c) investigated the effect of Triton X-100 added at the binding stage (2 hours ice incubation, 1 hour standing on bench incubation, and then spinning followed by detergent addition before column loading) and reported only a small decrease of 7.63% in the yield compared to the control. Visual observations of clear solutions with no turbidity indicated that Triton X-100 appears to stop completely the aggregation process. This detergent reduces the association between the refolding intermediates to prevent correct folding (as well as aggregation) and this explains the zero yields obtained in the present study. Other workers who have used Triton X-100 to suppress aggregation have afterwards removed the detergent using compounds like cyclodextrin (Rozema and Gellman, 1996) before continuing with the next stages of purification. In the present work, because the effect of Triton X-100 on binding of Con A to the dextran

affinity matrix was relatively small (7.63 % decrease in yield), it was decided not to attempt removal in subsequent folding experiments.

A further study was conducted focusing on the timings of Triton X-100 additions (Table 3.8), to explore more deeply the refolding process concerning current ideas on Con A folding. The study investigated the refolding process under different conditions while using 0.27 and 0.5M Gnd-HCl final concentrations. (The resultant purified yields are of high quality especially when using 0.5M Gnd-HCL and this can be seen from (A_{280}/A_{260}) ratios.) The design of the experiments was planned to focus on the time-course of the refolding progress because results from the first study (Table 3.7) suggested that Triton X-100 could be used as a "tool" to stop folding completely at set intervals. Because it has been shown that there is then little effect of detergent on binding of already folded Con A to dextran, this allows measurements to be made of how much folding has taken place up to the moment of detergent addition.

The first set of experiments (a), using 0.5M Gnd-HCL final concentration and detergent addition at: 0 hour showed a 100% decrease relative to the control, at ½ an hour refolding on ice, indicated a vast decrease of 86.01%, at 1 hour on ice indicated a decrease of 74.09% , and at 2 hours on ice showed a decrease of 55.51%.. Triton X-100 addition at 4 hours (2 hours on ice, 2 hours on bench) indicated a decrease of 24.21% relative to the control. (b) The second set of experiments, using 0.27M Gnd-HCL final concentration and detergent addition at: 0 hour indicated 100% decrease relative to control, at 1 hour refolding on ice indicated a decrease of 45.96%, at 2 hours refolding (1 hour on ice, 1 hour on bench) indicated the same yield as the control i.e. 0% decrease. These experiments using 0.5M Gnd-HCl, (Table 3.8 (a), and Figure3.21) indicated a slower pace of folding compared to 0.27M Gnd-HCl (Table 3.8 (b), and Figure 3.22). However, fewer data points were obtained for the latter conditions.

An experiment (Table3.8 (a), and Figure 3.21) using "extreme" time periods was conducted to assess the time limits during the refolding process. This experiment was all night ice incubation and then all day standing at bench temperature and indicated a decrease of 13.92% compared to the control. This "extreme" experiment has shown that there are limits for the refolding process to occur and that after longer time at 0°C aggregation may still reduce the overall yields.

The most important findings are shown by a time-course plot of the more extensive data set (Figure 3.23), which showed a nearly constant rate of increase in the refolding yield obtained. On further investigation, this approach might give insights into the kinetics of

folding. It is significant that Chatterjee and Mandal (2005) have studied the quaternary association and reactivation of dimeric Con A (Figure 3.24) by refolding pure plant-derived lectin at pH 5.2, where the dimer rather than tetramer predominates. Their experiments show a low level of activity of the refolding protein up to 80 minutes followed by a faster rate of folding up to 200 minutes. It is striking that although recombinant Con A has been refolded here (Figure 3.23) under different conditions, nevertheless similar times are taken to recover similar levels of activity. Why the recombinant protein shows no lag in folding is not clear, but may be due to a number of factors: the 4.5-fold higher Con A protein concentration, the higher pH = 7.0 (which favours tetramer formation after dimerisation), lower temperature for most of the experiment, different buffer composition, or other factors.

An important correlation was found by Chatterjee and Mandal (2005) between regain of activity on refolding (Figure 3.24) and re-association of subunits to form dimers, which was determined in a separate experiment. That these two processes run in parallel strongly suggests that dimer formation is an essential early step in successful folding. The findings in the present work on refolding of this recombinant lectin are therefore in broad agreement with those of other workers on the plant-derived protein.

Timing of Triton X-100 Additions	Refolding Yield 0.5 M Gnd-HCl			Binding Yield 0.5 M Gnd-HCl		
	(mg/l culture)	A_{280}/A_{260}	Relative Yield %	(mg/l culture)	A_{280}/A_{260}	Relative Yield %
Control (No addition of Triton X-100)	23.3 26.7 M=24.9	2.0 1.9	100	*23.1* *26.7* *M=24.9*	*2.0* *1.9*	*100*
(a)TritonX-100 Addition at the start	0.0	___	0.0	____	____	___
(b)TritonX-100 addition after 2 hours ice incubation	12.1	4.5	48.6	____	____	____
(c)TritonX-100 addition before binding to chromatography columns (2hours on ice, 1 hour on bench, then spin, then add Triton X-100)	___	___	___	23.9 22.1 M=23	1.823 1.784	92.4

Table 3.7 Effect of the timings of Triton X-100 additions on the refolding and binding of Con A (At 0.5 M Gnd-HCl concentration)

Control sets with no additions of Triton X-100, 1hour ice incubation and 1hour standing on bench. Refolding variables were the addition of Triton X-100 (a) at the start, (b) after 2 hours ice incubation, and (c) after 1 hour standing on bench incubation. Binding conditions were the additions before binding to columns i.e. after 2 hours ice incubation, 1 hour standing on bench incubation, and then spinning. Table shows relative percentage yields compared to controls.

Refolding conditions	Timings of Triton X-100 Additions	Yield (mg/l culture)	(A_{280}/A_{260})	Relative Yield %
(a) Refolding 0.5M Gnd-HCl 3 hours: 2 hours on ice, 1 hour on bench	Control (No addition of Triton X-100)	37.4 29.1 M=33.2	4.2 3.1	100
	0 hour	0.0	-	0.0
	½ hour On ice	4.7	1.8	14.2
	1 hour On ice	8.6	4.9	26.0
	2 hours On ice	16.3 13.3 M=14.8	4.6	44.6
Refolding 0.5M Gnd-HCl 4 hours: 2 hours on ice, 2 hour on bench Incubation	4 hours: 2hours on ice, 2 hours on bench	25.2	1.8	75.9
Refolding 0.5M Gnd-HCl (extreme): All night on ice, All day on bench incubation	Extreme: All night on ice, All day on bench	28.6	2.8	86.1
(b) Refolding 0.27M Gnd-HCl 2 hours: 1 hour on ice, 1 hour on bench Incubation	Control (No addition of Triton X-100)	22.3	2.2	100
	0 hour On ice	0.0	-	0.0
	1 hour On ice	12.1	2.0	54.3
	2 hours 1 hour on ice, 1 hour on bench	22.3	2.0	100

Table 3.8 Effect of the timing of Triton X-100 additions on the refolding of Con A (At 0.5 and 0.27M Gnd-HCl Concentrations)

Shows the use of Triton X-100 as a tool to investigate time-course of refolding at 0.27 and 0.5 M Gnd-HCl final concentrations. Control sets are the refolding conditions with no additions of Triton X-100. Table indicates the high yield obtained at the control level when refolding at 0.5M and also during the 2 hours ice incubation period. Furthermore, the Table also indicates clearly the nearly double amount of the yield when doubling the time at which Triton X-100 was added.

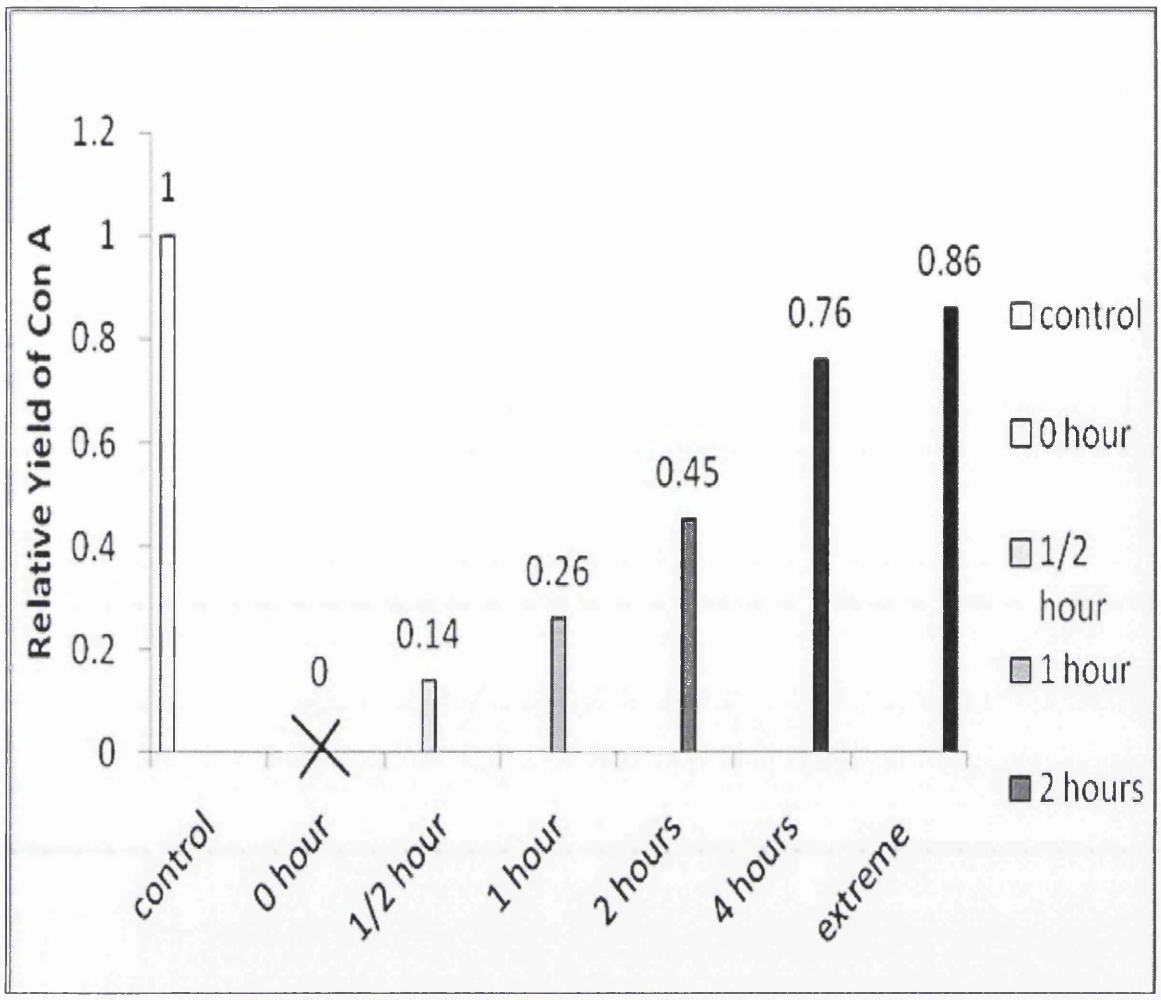


Figure 3.21 Effect of timings of TritonX-100 additions on the relative yields of Con A (when refolding at 0.5 M Gnd-HCl concentration)

Shows the relationship between the timings of additions of Triton X-100 and the relative yields. Control sets are the refolding conditions with no additions of TritonX-100. Figure indicates the increase of refolding yields with the increase in time. Figure used data from Table 3.8, part (a).

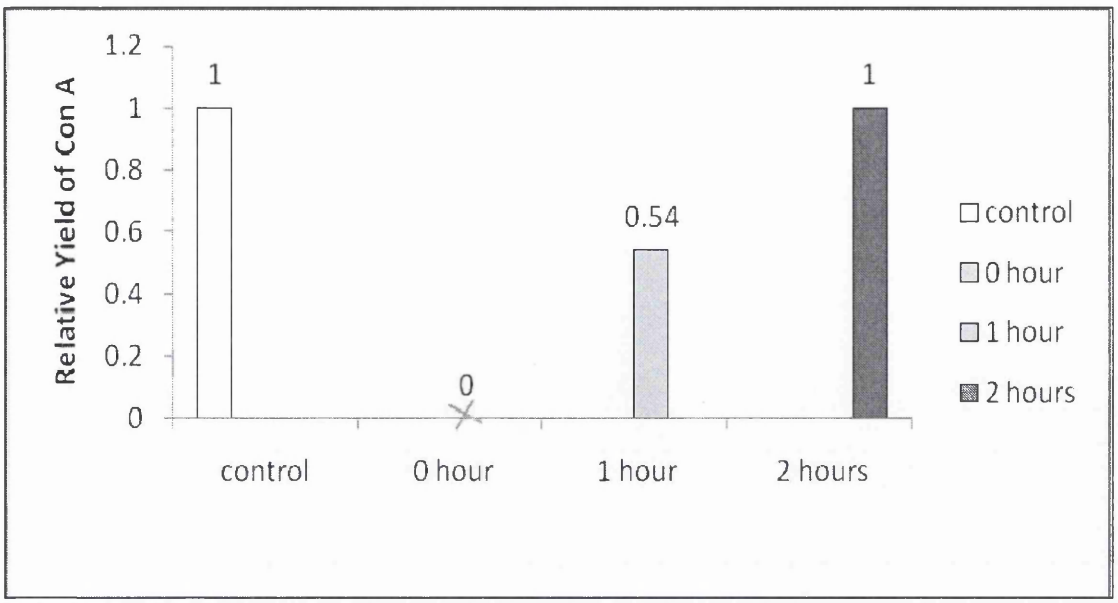


Figure 3.22 Effect of timings of TritonX-100 additions on the relative yields of Con A (when refolding at 0.27 M Gnd-HCl concentration)

Relates the yields to the timings of TritonX-100 additions at 0.27 M Gnd-HCl concentration. Control sets are the refolding conditions with no additions of TritonX-100. Refolded yields increased with increasing time in the absence of detergent.
Figure used data from Table 3.8, part (b).

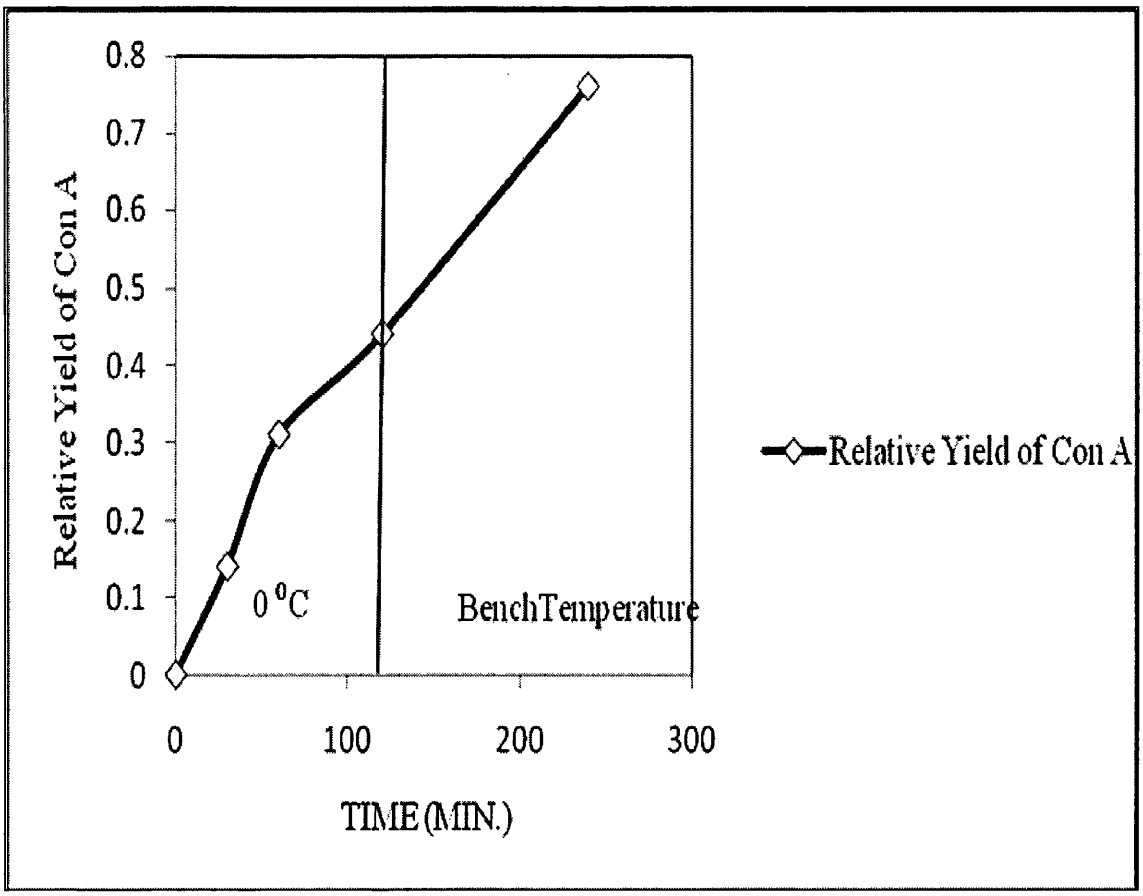


Figure 3.23 Time-course of folding of Con A

Shows the variation in relative yield with the time of the addition of Triton X-100. Detergent was added at time intervals of 0, 30, 60, 120, and 240 minutes. Yield is shown relative to the Control {33.2 mg/l obtained after refolding (without Triton X-100 addition) over 3 hours: 2 hours ice incubation, and then 1 hour standing on bench, as in Table 3.8. The concentration of Con A available for refolding is estimated from the control yield as 92 μg/ml or 3.6 μM (with respect to monomer).}. Figure indicates the time-course of the refolding process for recombinant Con A at pH 7.0 (in a crude mixture of bacterial inclusion body material) and shows how with time the refolding yields steadily increase. Based on data from Table 3.8, part (a).

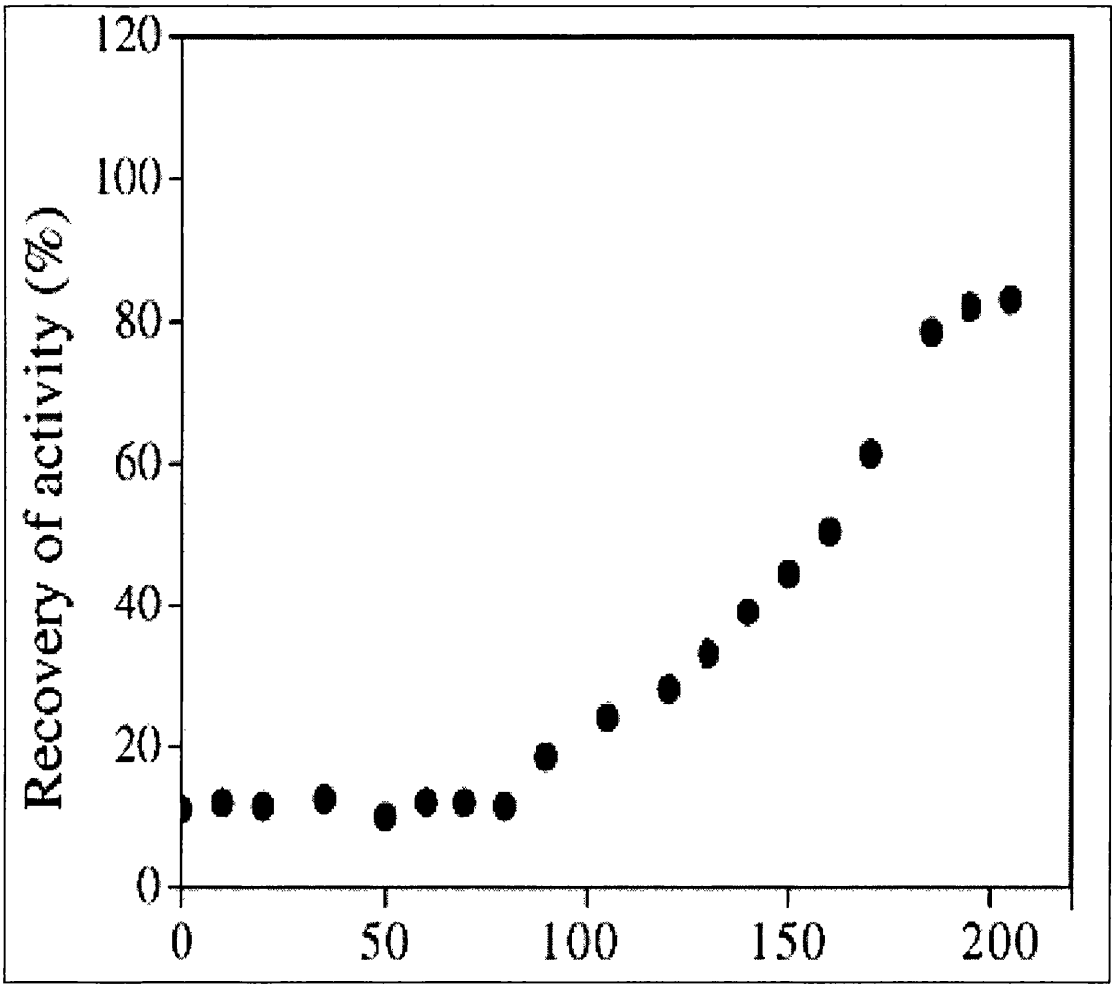


Figure 3.24 A plot of percent reactivation against time for pure plant Con A

Shows a plot of percent reactivation against time for pure plant Con A refolding at pH 5.2 as determined by a combination of affinity technique and OPA fluorescent method of protein estimation. The data are shown for a protein concentration of 0.8 μM and a constant temperature of 25°C throughout. Reproduced from **Chatterjee, A., Mandal D.K., (2005).** *International Journal of Biological Macromolecules 35: 103–109.*

Chapter Four

General Discussion and Conclusion

4.1 Introduction

The current project tried to optimise the refolding conditions of recombinant Con A by preventing or reducing the aggregation process without detriment to the folding processes producing active soluble lectin. The five different studies conducted as part of the project found that aggregation is a major cause of poor refolding yields that are the main barrier encountered during refolding in industrial settings.

4.2 Glycerol

The study investigated the effect of different glycerol concentrations on the refolding and binding stages of Con A. Results revealed that glycerol not only negatively affects the over all renaturation yields but also has a greater impact on the binding behaviour of the lectin towards the dextran (Table 3.1 and Figure 3.3). This result is consistent with a study conducted by Lopez-Jaramillo, et al., (2004) who observed that glycerol can occupy the saccharide binding sites in Con A crystals (Figure 3.6). This provides an explanation for the interference of glycerol with binding of Con A to the dextran affinity matrix observed here. Glycerol is widely celebrated as a refolding assistant with other proteins Meng, et al., (2001). However, the current study excludes its effects on the refolding of Con A as proteins differ in their actions towards different refolding conditions. Therefore, as a conclusion, it is assumed that glycerol competes with dextran for the carbohydrate-binding site of Con A so reducing recovery yields. However, glycerol is not very effective in eluting Con A which has already bound to dextran in its absence, compared with elution by methyl-∝-D-mannopyranoside (the saccharide normally used for specific elution). Glycerol was not investigated further as a folding aid for Con A due to these specific binding problems reducing recovery from the affinity columns.

4.3 Guanidine Hydrochloride (Gnd-HCl)

The study explored the effect of different low concentrations of guanidine hydrochloride on the refolding and binding of Con A as reported in literature (refer to section 1.4.7.1). The present study ruled out the use of concentrations of 0.75 to 1.5M as unsuitable since their effect on the binding and refolding of Con A produced an obvious decrease on the renaturation yields (Table 3.2, Figures 3.8 and 3.9). Moreover, the comparison between the 0.27 and 0.5 M Gnd-HCl concentrations (Table 3.2) during the phases of refolding and binding indicated the more positive effect of 0.5 M concentration on the refolding of Con A than 0.27M. Although, 0.27 M Gnd-HCl produces acceptable refolding yields, 0.5 M Gnd-HCl concentration appeared to improve the overall yields. The use of Gnd-HCL at 8M completely unfolds Con A, while diluting to 0.27 M allows for its refolding. Raising the concentration to 0.5M allows for more refolding since it can prevent aggregation to some extent. This may be through the denaturant increasing the solubility of a hypothetical unstable monomer. A review by De Bernardez Clark (2001) pointed out that, although low Gnd-HCl concentration prevent aggregation the use of concentrations less than 0.25 M inhibited the dimerization of the platelet-derived growth factor (PDGF). The present findings have opened a new door to obtain higher refolding yields by adopting the 0.5M final Gnd-HCl concentration during refolding in further studies.

4.4 Dilution

The third study investigated the role played by different dilution protocols on the production of biologically active Con A. In this study rapid, continuous, pulse and step-wise dilution are explored. The first part exploited different dilution factors of the rapid protocol to see their effect combined with 0.27 and 0.5 M Gnd-HCl on the refolding yield (Table 3.3). The dilution factor 30X indicated the maximum yield obtained using 0.5 M Gnd-HCl concentration. Further, the different protocols were compared to controls (Table 3.4 and Figure 3.12) where the continuous dilution protocol has shown the best yield by a small margin obtained when using the 0.5 M Gnd-HCl concentration compared with the other protocols. However, manual continuous dilution is much more laborious in practice than rapid dilution. The small margin of improvement did not

justify further work to automate continuous dilution using pumped systems. The study has also shown that step-wise dilution is an un-desirable method for the refolding of Con A as it reduces the overall yields by the forming large amounts of aggregates. Moreover, the study relates dilution 30X (protein concentration) with Gnd-HCl concentration (0.5 M), identified as appropriate from the previous study, as suitable factors in producing optimum yields when combined together. This work initiated examination of the relationship between the protein concentration and the renaturation yield (Figure 3.16). It is seen that high protein concentration gives rise to aggregation, while lowering the protein concentration produces increasing yields. However, when protein concentration is lowered further the yields decrease again. This observation is explained by a hypothetical scheme (Figure 3.15) indicating the formation of a fully folded dimer through the formation of a partially folded monomer. Two possible pathways are depicted: Association 1 is less sensitive to dilution as stable monomer exists with unlimited time available to find a partner and proceed to form a fully folded dimer and complete the folding process. {This is represented by the expected line in (Figure 3.16).} Association 2, where unstable partially folded monomer exists, is more sensitive to dilution with a limited time to fold. This unstable partially folded monomer might follow different pathways: either aggregating at high protein concentrations, or forming an inactive soluble monomer, or succeeding in following the Association2 pathway to find a partner and form a partially folded dimer. {This represented by the actual line in (Figure 3.16).} These hypotheses are in agreement with the general principles reviewed by Price (1994) when considering the re-folding of multi-subunit proteins. High protein concentration favours aggregation that is produced by non-specific hydrophobic surfaces exposed early in the refolding pathway. Low enough concentrations of protein cause the rate of association reactions to become so slow that non-associated species will exist for significant periods of time. These intermediates are of limited stability so may be degraded by proteases, chemically modified or adsorbed onto surfaces such as the walls of the reaction flask. This explains why the yield of active multi-meric protein declines at low concentrations. As protein concentration is varied, these competing trends of aggregation versus "correct" folding/association give rise to curves showing a maximum yield of active protein at an optimum concentration. An example of such a curve of concentration versus reactivation by Rudolph et al. (1977) is shown in (Figure 4.1) and shows a maximum. This resembles the plot given earlier (Figure 3.16) and indicates that the fate of an unstable intermediate is critical for

correct folding. To conclude, this dilution study is very important in understanding the pathway followed by Con A to ensure correct refolding. Experimentally determined limits to dilution enable optimised refolding yields.

4.5 Temperature

The study concerned the effect of different incubation temperatures on refolding through tackling the aggregation process. Usually lower temperatures reduce aggregation since hydrophobic interactions are weakened. Temperature is of vital importance and optimising it might lead to further improvement of renaturation yields. Table (3.6a), indicated that a very low temperature (- 5.5 °C) does not aid the refolding process and that at very high temperatures aggregation occurs with no detectable yields. Although continuous incubation at 0°C produced the highest yields, problems with column blockages were a disadvantage of refolding at this temperature. In contrast to this, a gradual warming up step produced some improvement in yield especially when incubating initially at 0°C. The study agrees with Yoshii et al. (2000) in that low temperatures have a definite role in the prevention of hydrophobic aggregation where an improvement of yield occurred. The results are also in agreement with the techniques of Wetlaufer and Xie (1996). Moreover, this study initiated use of a 2 hour ice incubation for the first time in our laboratory and this proved to be an important development in achieving greater yields since it allowed more time for the refolding process to complete. The gradual warming up step after a longer incubation on ice is another achievement that improved the overall yield. So to conclude, this was a good step towards optimising the refolding process and therefore was incorporated in the subsequent study.

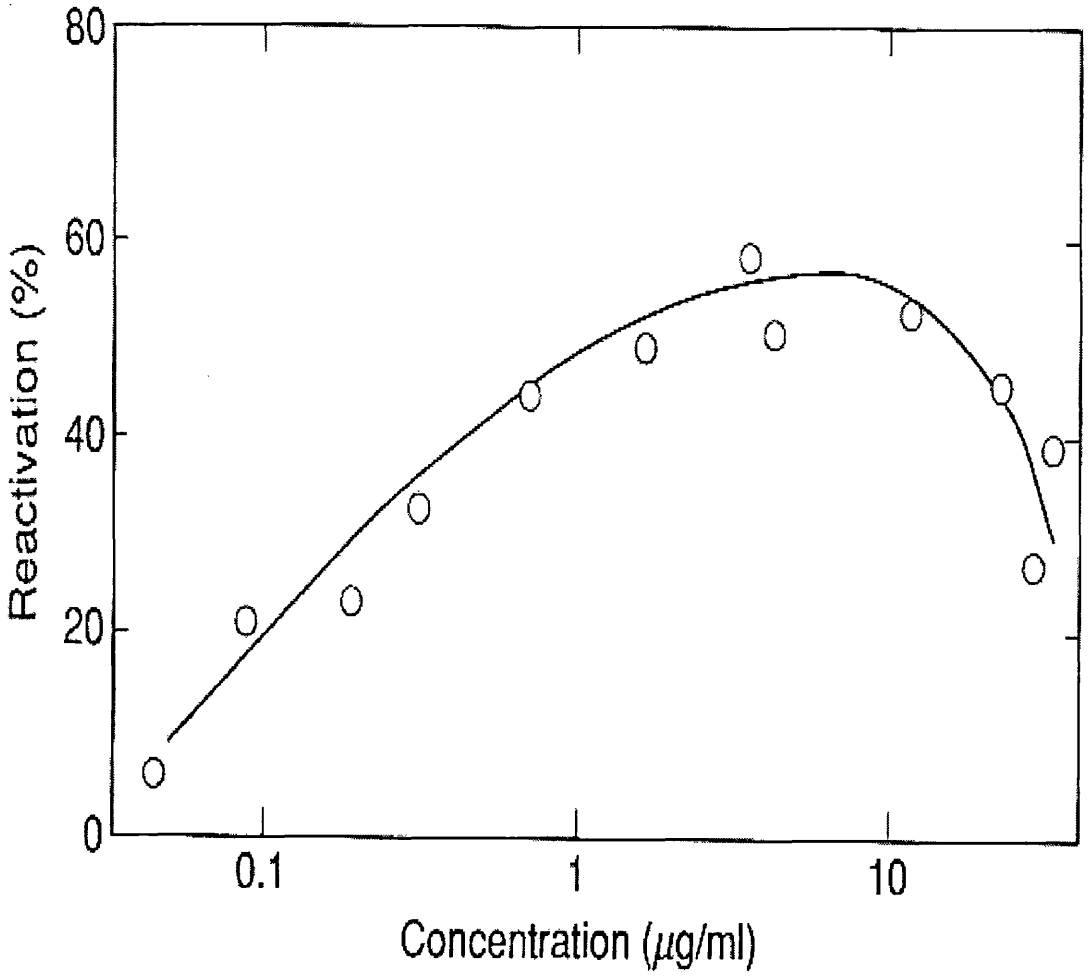


Figure 4.1 The regain of activity of lactate dehydrogenase (H4 isoenzyme)

Shows the after denaturation of lactate dehydrogenase (H_4 isoenzyme) as a function of protein concentration. A number of different methods of denaturation (low pH, Gdn-HCl, urea) were used in these experiments. Reproduced from (**Rudolph et al., 1977).** Biochemistry, **16**: 3384.

4.6 TRITON X-100

The fifth study explored different aspects of Triton X-100, a non-ionic detergent that is extensively used in the washing and solubilisation of inclusion bodies. The effect of Triton X-100 was first investigated on the refolding and binding of Con A (Table 3.7), where a zero yield was obtained at the refolding stage but only a small decrease was found in binding. This distinct result indicated that Triton X-100 interacts with the refolding intermediates at an early stage of the refolding process while at the same time its effect on binding to the dextran is very limited. These findings were developed into the idea of using Triton X-100 as a method to explore further the refolding process through the adoption of certain timings of Triton X-100 additions to the refolding mixtures, knowing that no effect will occur on the binding to the dextran stage. Different timings were selected, Gnd-HCl concentrations of 0.27 and 0.5 M, different incubation periods on ice and on bench, and all with the 30X factor of the rapid dilution protocol (relating the previous studies with the current one).

A plot of the relative yield versus time (Figure 3.23) resulted from this study and indicated the progress of the refolding process. This agreed to some extent with a plot by Chatterjee and Mandal (2005) (Figure 3.24). Moreover, their study also related to the quaternary association and reactivation of dimeric Con A. Furthermore, these findings were in accordance with the hypothetical Association 2 pathway (Figure 3.15) and the actual plot (Figure 3.16) shown earlier in the dilution study. This novel demonstration of the great sensitivity of Con A folding to inhibition by Triton X-100 shows that an essential step on the correct folding pathway depends on hydrophobic interactions taking place. It seems likely that these interactions are between hydrophobic surfaces essential for association of the subunits so that they can form dimers and complete their folding. Detergent inhibition of folding is an easy and inexpensive method which can be developed to investigate further aspects of the refolding of proteins. Furthermore, detailed experiments are required and the use of purified proteins with accurately known concentration values is desirable in this novel technique.

4.7 Conclusions

As reviewed in the Introduction (Section 1.3), aggregation is the major problem facing large scale production and maximum yield. Aggregation occurs when intra- and intermolecular interactions driving protein folding are not correctly balanced during the renaturation or refolding process. Major sources of aggregation are: hydrophobic interactions, unfavourable apolar-polar interactions, high protein concentration, high concentrations of refolding intermediates, and non-specific aggregation.

The current project tried to avoid aggregation and improve the overall protein yield where it identified glycerol as an unsuitable aid in the refolding of Con A due to the binding effect it possessed towards the affinity dextran. Further results recommended the use of 0.5 M Gnd-HCl concentration as the initial point that guarantees increased refolding yields. The rapid and continuous dilution protocols are more suitable than the other exploited refolding techniques used in this project in improving the refolding yields of Con A. However, rapid dilution is still the preferable method to use since it is simple and reproducible. The 30X dilution factor has been demonstrated to be the optimum to enhance the refolding yields. A gradual warm-up step ensured the recovery of most of the refolded Con A and a 0°C temperature incubation combined with this technique improved the overall yield to maximum levels. Lengthening the ice incubation period to 2 hours instead of 1 hour allowed for more yield to be recovered. Triton X-100 does not aid the refolding of Con A due to its inhibitory effects on essential hydrophobic associations. However, its use as a tool is considered a breakthrough for estimating the time-course of refolding of the lectin, providing that more detailed experiments are conducted to improve the method. The combined findings allowed for development of a hypothetical scheme (Figure 3.15) to help understand the refolding pathway of Con A.

The major findings for increasing yield in this work can all be explained by their effects of decreasing the likelihood of aggregation of partially-folded intermediate forms of Con A. This is in general agreement with the ideas widely reported for other proteins and reviewed in the literature (as explained in the Introduction). In addition, Triton X-100 was found to suppress desired folding interactions in addition to any unwanted aggregation, so that the yield of active Con A was reduced to zero in its presence. However, this initially disappointing result became an unexpected bonus when it was shown that Triton X-100 could be used to investigate the time course of Con A folding.

Because this detergent was shown not to interfere with dextran binding by already folded lectin, it could be used to stop further folding occurring in a mixture, so that the amount of already folded protein could be easily measured at any desired time point. These experiments demonstrated that the time course of refolding determined in the present work for recombinant Con A was in accord with that for native Con A from jack beans (Section 3.1.4.3 and Section 4.4). The effects of Triton X-100 point to the importance of hydrophobic interactions at subunit interfaces which are essential for assembly of dimers of biologically active Con A.

My combined strategy improved the renaturation yield from only 12 mg/l culture to 33 mg/l culture i.e. a 2.75-fold increase. Other studies (Dong et al., 2007; Wang et al., 2008) following different dilution protocols have also reported an improvement in their final yields, however, these enhancements were not as high as I achieved in this project.

New general protocols have been developed for the refolding of proteins. The REFOLD (2005) database (http://refold.med.monash.edu.au) provides over 200 new protocols for overexpression, solubilisation and refolding of recombinant proteins. Today, protein refolding screening kits are available on the market: (www.moleculardimensions.com) and also the iFOLD protein refolding system from Novagen (www.novagen.com/ifold). These are designed to ensure accuracy and facilitate the tedious and time consuming refolding process, while minimising the budget needed for large volumes of buffers and refolding vessels.

Finally, from what has been achieved in this project, a wider approach to refolding can be followed in future studies utilising the continuous dilution protocol with an automated device that can control the volume of the denatured material added at an appropriate rate to refolding buffers. By using a stirred thermostated vessel temperature ranges can be changed as required. This would be advantageous for scaling up the process for industrial applications.

References

Agrawal, B. B. L., and Goldstein, I. J. (1972). Concanavalin A, the jack bean (Canavalia ensiformis) Phytohemagglutinin. Methods in Enzymology. **28:** 314.

Andrews, B. , Adari, H.,Hannig, G., Lahue, E.,Gosselin, M.,Martin, S., Ahmed, A., Ford, P. J., Hayman, Makrides, S. C., (1996). A tightly regulated high level expression vector that utilizes a thermosensitive *lac* repressor: production of the human T cell receptor Vβ5.3 in *Escherichia coli* Gene **182**: 101-109.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. Science **181:** 223 Ann. N. Y. Acad. Sci., **234**: 283-296

Ballery, N., Desmadril, M., Minard, P., and Yon, J. M. (1993). Biochemistry 32: 669.

Baneyx F. (1999). Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol, **10**:411-421.

Batas, B., Schiraldi, C., Chaudhuri, J.B. (1999). Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. J. Biotechnol. **68**: 149–15.

Becker, J. W., Cunningham G., Reeke, G. N. Jr., Wang, J. L., and Edelman, G. M. (1976). The Molecular Structure of Concanavalin A. In: Bittiger, H., and Schnebli, H. P. (Eds) Concanavalin A as a tool, Wiley and Sons pp 33-54.

Bittiger, H., and Schnebli, H. P. (Eds) Concanavalin A as a tool, Wiley and Sons pp 33-54.

Bowden, G.A., Georgiou, G. (1990). Folding and aggregation of _α-lactamase in the periplasmic space of Escherichia coli. J. Biol. Chem. **265:** 16760–16766.

Bowles, D. J, Marcus, S., Pappin, D. J. C., Findlay, J. B., Eliopoulos, E., Maycox P. R., Burgess, J. (1986). Posttranslational processing of Concanavalin A precursors in jack bean cotyledons. Journal of Cell Biology **102:** 1284–1297

Brosius, J., Erfle, M., and Storella, j., (1985). Spacing of the -10 and -35 regions in the tac promoter. Effect on its in vivo activity. J. Biol. Chem., **260**: 3539-3541.

Carrio, M. M., and Villaverde, A. (2002). Construction and deconstruction of bacterial inclusion bodies. Journal of Biotechnology. **96**: 3-12.

Chaozhan Wang, Lili Wang, Xindu Geng, (2008). Optimization of refolding with simultaneous purification of recombinant human granulocyte colony-stimulating factor

from *Escherichia coli* by immobilized metal ion affinity chromatography.Biochemical Engineering Journal, **In Press, Corrected Proof.**

Charron, C., Kadri, A., Robert, M. C., GiegeÂ, R. & Lorber, B. (2002). Crystallization in the presence of glycerol displaces water molecules in the structure of thaumatinActa Cryst. D58, 2060±2065.

Chatterjee, A. and Mandal, D. K. (2003). Denaturant-induced equilibrium unfolding of Concanavalin A is expressed by a three-state mechanism and provides an estimate of its protein stability. Biochim. Biophys. Acta. **1648:** 174-183.

Chatterjee, D.K. Mandal. (2005). Quaternary association and reactivation of dimeric concanavalin A International Journal of Biological Macromolecules **35:** 103–109.

Chrispeels, M. J., Hartl, P. M., Sturm, A., Faye, L. (1986). Characterization of the endoplasmic reticulum-associated precursor of Concanavalin A. Journal of Biological Chemistry **261:** 10021–10024.

Clark, E. D. B. (1998). Refolding of recombinant proteins. Current Opinion in Biotechnology. **9**: 157-163

Creighton TE (1997), How important is the molten **globule for correct** protein folding? Trends Biochem Sci 22:6-10.

Creighton, T. E. (1994). The protein folding problem. In Pain, R. H. (Ed) Mechanism of protein folding, IRL Press at Oxford University Press pp 1-25

Cunningham, B. A., Wang, J. L., Waxdal, M. J., and Edelman, G. M. (1975) The covalent and three-dimensional structure of Concanavalin A. II Amino acid sequence of cyanogens bromide fragment F_3. J. Biol. Chem. 1503-12

Dale, J., and Park, S. F., (2004). Molecular Genetics of Bacteria. 4th ED. Wily and Sons. ISBN 0 470 85084 1.

De Bernardez Clark E. (2001). Protein refolding for industrial processes. Curr Opin Biotechnol. **12:202**–207.

De Bernardez Clark, E., (1998). Refolding of recombinant proteins. Curr. Opin. Biotechnol. **9**:157–163.

De Bernardez Clark, E., Schwarz, E., Rudolph, R. (1999). Inhibition of aggregation side reactions during in vitro protein folding, Methods Enzymol. **309**: 217–236.

De Boer, Comstock, L. J., and Vasser, M., (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc. Nat.Acad. Sci. USA., **80**: 21-25

Deacon. (1997).J. Chem. Soc., Faraday Trans., **93 :**(24).

Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J., Papiz, M. Z., Edelman, G. M., Cunningham, B. A., Reeke, G. N. Jr., Becker, J. W., Waxdal, M. J. and Wang, J. L. (1972). The covalent and three-dimensional structure of Concanavalin A. Proc. Nat. Acad. Sci. **69:** 2580-4

Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J., Papiz, M. Z., Wan, T., and Campbell, J. (1989). The structure of saccharide-binding site of Concanavalin A. EMBO Journal. **8:** 2189-2193

Emmerich, C., Helliwell, J. R., Redshaw, M., Naismith, J. H., Harrop, S. J., Raftery, J., Kalb (Gilboa), A. J., Yariv, J., Dauter, Z., and Wilson, K. S. (1994). High-resolution structures of single-metal-substituted concanavalin A: The Co, Ca-protein at 1.6 Å and the Ni, Ca-protein at 2.0 Å, Acta Crystallogr. **D50:** 749–756.

Fahey EM, Chaudhuri JB, Binding P., (2000). Refolding of low molecular weight urokinase plasminogen activator by dilution and size exclusion chromatography — a comparative study. Sep Sci Technol, **35**:1743-1760.

Fan-Guo, M., Yong-Doo, P., Hai-Meng, Z., (2001). Role of proline, glycerol, and heparin as protein folding aids during refolding of rabbit muscle creatine kinase Int. J. of Bio. & Cell Biol. **33**, 701-709.

Farina, R. D., and Wilkins, R. G. (1980). Kinetics of interaction of some a- and b-D-monosaccharides with Concanavalin A, Biochim. Biophys. Acta **631:** 428–438.

Futami J, Tsushima Y, Tada H, Seno M, Yamada H., (2000). Convenient and efficient in vitro folding of disulfide containing globular protein from crude bacterial inclusion bodies. J Biochem, **127**:435-441.

Georgiou, G., Telford, J.N., Shuler, M.L., Wilson, D.B., (1986). Localization of inclusion bodies in Escherichia coli overproducing _-lactamase or alkaline phosphatase. Appl. Environ. Microbiol. **52**: 1157–1161.

Georgiou, G., Valax, P., (1999). Isolating inclusion bodies from bacteria. Methods Enzymol. **309**: 48–58.

Goeddel, D. V., and Gold, L. (1990). Systems for heterologous gene expression and expression of heterologous protein in E. coli In Goeddel, D V., (Ed) Gene Expression Technology. Methods in Enzymology **185:** 3-15 Academic Press.

Goldberg ME, Expert-Bezancon N, Vuillard L, Rabilloud T., (1996) Nondetergent sulfobetaines: a new class of molecules that facilitate protein renaturation. Fold Des, **1**:21-2

Goldenberg, D. P., (1989). Circularly permuted proteins. Protein Engineering, **2**: 493-495.

Goldstein, I. J. and Poretz, R. D. (1986). Isolation, Physicochemical Characterisation, and Carbohydrate-Binding Specificity of Lectins. In: **Liener, I. E., Sharon, I. J., and Goldstein (Eds) (1986).** The Lectins: Properties Functions and Applications in Biology and Medisine, Academic Press Inc Orlando pp. 33-248

Goldstein, I. J., Hollerman, C. E., and Smith, E. E.,(1965). Protein-carbohydrate interaction. II. inhibition studies on the interaction of Concanavalin A with polysaccharides. Biochemistry **4**: 876-883.

Goldstein, M. A., and Doi, R.H., (1995). Prokaryotic promoters in biotechnology Biotechnol. **6**: 491-493.

Gu Z, Weidenhaupt M, Ivanova N, Pavlov M, Xu B, Su Z-G, Janson JC., (2002). Chromatographic methods for the isolation of, and refolding of proteins from, Escherichia coli inclusion bodies. Protein Express Purif, **25:**174-179.

Hamodrakas, S.J., Kanellopoulos, P.N., Pavlou, K., Tucker, P.A., (1997) The crystal structure of the complex of Concanavalin A with 4'-methylumbelliferyl-alpha-D-glucopyranoside. journal of structural biology **118,** 23–30

Hannig, G., and Makrides, S.C. (1998). Strategies for optimising heterologous protein expression in Escherichia coli. Trends in Biotechnology. **16**: 54-60.

Hardman, K. D., Agarwal, R. C., and Freiser, M. J. (1982) Manganese and calcium binding sites of concanavalin A, J. Mol. Biol. **157:** 69–86.

Hartman JR, Mendelovitz S, Gorecki M., (1999). Refolding of proinsulins without addition of reducing agents. US patent, 6001604.

Hayashi, M. N., and Hayashi, M., (1985). Cloned DNA sequences that determine mRNA stability of bacteriophage phi X174 in vivo are functional. Nucleic Acids Res., **13**: 5937-5948.

Hohenblum, H., Vorauer-Uhl, K., Katinger, H., Mattanovich, D., (2004). Bacterial expression and refolding of human trypsinogen. J. Biotechnol. **109**: 3–11.

Huxtable S, Zhou H, Wong S, Li N., (1998). Renaturation of 1-aminocyclopropane-1-carboxylate synthase expressed in Escherichia coli in the form of inclusion bodies into a dimeric and catalytically active enzyme. Protein Express Purif, **12**:305-314.

Jaenicke, R., and Rudolph, R. (1989). Folding proteins. In Creighton, T. E (Ed) Protein Structure a practical approach, IRL Press at Oxford University Press pp 191-223

Jain, D., Kaur, K. J. & Salunke, D. M. (2001a). Structural Basis of Functional Mimicry between Carbohydrate and Peptide Ligands of Con A. Biophys. J. **80**: 2912± 2921.

Jones, D. H., (1995). Folding, activation and protein splicing of recombinant Concanavalin A precursors: an exceptional protein to prove some rules. Perspectives on Protein Engineering & Complementary Technologies. Geisow, M. and Epton, R. (eds.) Mayflower Worldwide, Birmingham pp. 70-73.

Jones, D. H., (2005) Purification method for a recombinant glucose binding protein. WO20051987.Patent.

Jungbauer, A., Kaar, W., Schlegl, R., (2004). Folding and refolding of proteins in chromatographic beds. Curr. Opin. Biotechnol. **15**: 487–494.

Jungbeuer, A., and Kaar, W., (2007). Current status of technical protein refolding. Journal of Biotechnology 128: 587-596.

Kanellopoulos, P. N., Pavlou, K., Perrakis, A., Agianan, B., Vorgias C. E., Mavrommatis, C., Soufi, M., Tucher, P. A., Hamodrakas, S. J. (1996). The crystal structure of the complexes of Concanavalin A with 4'-Nitrophenyl-α-D-mannopyranoside and 4'-Nitrophenyl-α-D-glucopyranoside. Journal of Structural Biology **116:** 345-355

Karplus, M. (1997). The Levinthal paradox: yesterday and today. Folding & Design **2:** S69-S75.

Katoh, S. and Katoh, Y. (2000). Continuous refolding of lysozyme with fed-batch addition of denatured protein solution Process Biochem.**35**: P. 1119-1124.

Katzav-Gozansky T., Hanan E., Solomon B. (1996). Effect of monoclonal antibodies in preventing carboxypeptidase A aggregation. Biotechnol. Appl. Biochem. **23**: 227.

Khan, R.H., Rao, K.B., Eshwari, A.N., Totey, S.M., Panda, A.K. (1998). Solubilisation of recombinant ovine growth hormone with retention of native-like secondary structure and its refolding from the inclusion bodies of Escherichia coli. Biotechnol. Prog. **14**: 722–728.

Khodagholi, F.R. Yazdanparast F., (2005) Artificial chaperone-assisted refolding of chemically denatured α-amylase. International Journal of Biological Macromolecules, 35:257-263

Kumagai, I., Tsumoto, K., (1998). Generation of novel functional antibody molecules by in vitro selection system, Tanpakushitsu Kakusan Koso **43:** 201–205.

Kumar TK, Samuel, D., Jayaraman, G., Srimathi, T., Yu, C., (1998) The role of proline in the prevention of aggregation during protein folding in vitro. Biochem. Mol. Biol. INT. **46:** 509-517.

Larson, E. L., and Mattiasson, B., (1994). Isolation of Concanavilin A by affinity precipitation. Biotechnology Tech., **8 (1)**: 51-56.

Lee Y, Lee H, Han K. (2000) Process for the preparation of active somatotropin from inclusion bodies. International application WO 00/12537.

Lee, C.T., Buswell, A.M., Middelberg, A.P.J., (2002). The influence of mixing on lysozyme renaturation during refolding in an oscillatory flow and a stirred-tank reactor. Chem. Eng. Sci. **57**: 1679–1684.

Lee, C.T., Mackley, M.R., Stonestreet, P., Middelberg, A.P.J., (2001). Protein refolding in an oscillatory flow reactor. Biotechnol. Lett. **23**:1899–1901.

Levinthal, C. (1968). Are there pathways for protein folding? J. Chim. Phys. **65:** 44

Lilie, H., Schwarz, E., Rudolph, R., (1998). Advances in refolding of proteins produced in E. coli. Curr. Opin. Biotechnol. **9**: 497–501.

LoÂpez-Jaramillo, F. J., GonzaÂlez-RamõÂrez, L. A., Albert, A., Santoyo-GonzaÂlez, F., Vargas-Berengueld, A., and OtaÂloraa F., Acta Cryst. (2004). Structure of concanavalin A at pH 8: bound solvent and crystal contacts D60 1048±1056

Maachupalli-Reddy J, Kelley BD, De Bernardez Clark E., (1997). Effect of inclusion body contaminants on the oxidative renaturation of hen egg white lysozyme. Biotechno/Prog, **13**:144-150. 125–135

Makrides SC., (1996). Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol Rev, **60:**512-538.

Meng, F., Park, Y., Zhou, H., (2001). Role of proline, glycerol, and heparin in protein folding aids during refolding of rabbit muscle creatine kinase, Int. J. Biochem. Cell Biol. **33**: 701–709

Middelberg, A. P. J. (2002). Preparative protein refolding. Trends in Biotechnology. **20**: 437-443

Min W. (1992). Emulation of post-translational processing of Concanavalin A by recombinant DNA manipulations. University of Wales Swansea, PhD thesis.

Min, W. and Jones, D. H., (1994). In vitro splicing of Concanavalin A is catalysed by asparaginyl endopeptidase Nature Struct. Biol., **1:** 502-504

Mitraki, A., and King, J. (1989). Protein folding intermediates and inclusion body formation Biotechnology **7:** 690-697

Mitraki, A., King, J., (1992). Amino acid substitutions influencing intracellular protein folding pathways.FEBS Lett. **307**, 20-25.

Muller C, Rinas U., (1999). Renaturation of heterodimeric platelet-derived growth factor from inclusion bodies of recombinant Escherichia coli using size-exclusion chromatography. J Chromatogr A, **855**:203-213.

Naismith, J. H., and Field, R. A. (1996). Structural basis of CON A–a-MUG STRUCTURE 29 trimannoside recognition by concanavalin A, J. Biol. Chem. **271:** 972–976.

Naismith, J. H., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Wan, T. C. M., Weisgerber, S., Kalb(Gilboa), A. J., and Yariv, J. (1993). Refined structure of cadmium-substituted concanavalinAat 2.0 Å resolution, Acta Crystallogr. **D49:** 561–571.

Naismith, J.H., Emmerich, C., Habash, J., Harrop, S.J., Helliwell, J.R., Hunter, W.N., Raftery, J., Kalb (Gilboa), A.J. and Yariv, J. (1994). Refined structure of concanavalin A complexed with methyl alpha-D-mannopyranoside at 2.0 A resolution and comparison with the saccharide-free structure. Acta Cryst. **50**: 847 – 858

Nishihara, K., Kanemori, M., Yanagi, H., Yura, T., (2000). Overexpression of trigger factor prevents aggregation of recombinant proteins in Escherichia coli. Appl. Environ. Microbiol. **66**, 884–889.

Nishihara, T., Iwabuchi, T., and Nohno, T., (1994). A T7 promoter vector with a transcriptional terminator for stringent expression of foreign genes. Gene, **145**: 145-146.

Novagen pET system Tutorial. (2002-2003 Catalogue).
http://www.novagen.com/SharedImages/Novagen/05_PROEXP.pdf assessed August 2003.

Novagen Technical Bulletin (1993), 71.

Novy, R. and Morris, B. (2001). Use of glucose to control basal expression in the pET system, in**Nova**tions **13**, 8–10.

Pace, C. N., Shirley, B. A., Thomson, J. A. (1989). Measuring the conformational stability of a protein. In Creighton, T. E (Ed) Protein Structure a practical approach, IRL Press at Oxford University Press pp 311-345

Pain RH., (1994). Mechanismsof protein folding. Oxford: **IRL** Press. Oxford

Patra, A.K., Mukhopadhyay, R., Mukhija, R., Krishnan, A., Garg, L.C., Panda, A.K., (2000). Optimization of inclusion body solubilisation and renaturation of recombinant human growth hormone from Escherichia coli. Protein Expr. Purif. **18**:182–192.

Peng Z, Kim PS., (1994). A protein dissection study of molten globule.. Biochemistry, **33**:2131-2141.

Poole, E. S., Brown, C. M., and Tate, W. P.(1995). Translational termination efficiency in both bacteria and mammals is regulated by the base following the stop codon. EMBO J.**14**: 151-158

Poretz, R. D. and Goldstein, I. J. (1970). An examination of the topoghraphy of the saccharide binding sites of Concanavalin A and of the forces involved in complexation. Biochemistry,9: 2890-6.

Price, N. C., (1994). Assembly of multi-subunit structures. Chapter **7**: 160-193. In Pain RH., (1994). Mechanismsof protein folding. Oxford: **IRL** Press. Oxford.

Ptitsyn OB, Bychkova VE, Uversky VN., (1995). Kinetic and equilibrium **folding intermediates.** Philos Trans R Soc Lend B Biol Sci, **348**:35-41.

Radford, S. (2000). Protein folding: progress made and promises ahead. Trends Biochem, Sci. **25:** 611-618

Ramis, C., Gomord, V., Lerouge, P., and Faye, L. (2001). Deglycosylation is necessary but not sufficient for activation of proconcanavalin A Journal of Experimental Botany **52:** 911-917

Rashid, F., Sharma, s., and Bano, B.,(2005). Comparison of Guanidine Hydrochloride (GdnHCl) and Urea denaturation on inavtivation and unfolding of human placental cystatin (HPC). The Protein Journal, **24**:283-292.

Rattenholl, A., Lilie, H., Grossmann, A., Stern, A., Schwarz, E., Rudolph, R., (2001). The pro-sequence facilitates folding of human nerve growth factor from Escherichia coli inclusion bodies. Eur. J. Biochem. **268**: 3296–3303.

Reddy A, Grimwood BG, Plummer TH, Tarentino AL., (1998). High-level expression of the endo-β-N-acetylglucosaminidase F2 gene in E. coli: one step purification to homogeneity. Glycobiology, **8**:633-636.

Redfield C, Smith RAG, Dobson CM., (1994). Structural characteriza- tion of a highly ordered 'molten globule' at low pH. Nature Struct Biol, **1**:23-29.

Reeke, G. N. Jr., Becker, J. W., and Edelman, G. M., (1975). The covalent and three-dimensional structure of Concanavalin A. IV Atomic coordinates, hydrogen bonding,

and quaternary structure. J. Biol. Chem. **250:** 1525-47

Rinas, U., Bailey, J.E., (1992). Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli. Appl. Microbiol. Biotechnol. **37**, 609–614.

Rosenberg, A.H., Lade, B.N., Chui, D.S., Lin, S.W., Dunn, J.J. and Studier, F.W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene, **56**: 125-135.

Rozema D, Gellman SH. (1996). Artificial chaperone-assisted refolding of denatured-reduced lysozyme: Modulation of the competition between renaturation and aggregation. Biochemistry, **35:**15760-15771.

Rozema, D., Gellman, S.H., (1996). Artificial chaperone-assisted refolding of carbonic anhydrase B. J. Biol. Chem. **271**:3478–3487.

Rudolph , R. Heider,I., Westhof, E., and Jaenicke, R. (1977). Mechanism of refolding and reactivation of lactic dehydrogenase from pig heart after dissociation in various solvent media. Biochemistry, **16**: 3384.

Rudolph, R., Fischer, S., (1990). Process for Obtaining Renatured Proteins. US patent **4**:933,494.

Schlegl, R., Iberer, G., Machold, C., Necina, R., Jungbauer, A., (2003). Continuous matrix-assisted refolding of proteins. J. Chromatogr. A **1009**: 119–132.

Schlegl, R., Necina, R., Jungbauer, A., (2005a). Continuous matrix assisted refolding of inclusion-body proteins: effect of recycling. Chem. Eng. Technol. **28**:1375–1386.

Schlegl, R., Tscheliessnig, A., Necina, R.,Wandl, R., Jungbauer, A.,(2005b). Refolding of proteins in a CSTR. Chem. Eng. Sci. **60**: 5770–5780.

Sheldon, P. S., Bowles, D. J. (1992). The glycoprotein precursor of Concanavalin A is converted to an active lectin by deglycosylation. EMBO Journal **11:** 1297–1301.

Sheldon, P. S., Keen, J. N., Bowles, D. J. (1998). Purification and characterization of N-glycanase, a Concanavalin A binding protein from jackbean (Canavalia ensiformis). Biochemistry Journal **330:** 13–20.

Shirley, B. A. **(1995)** Urea and Guanidine Hydrochloride Denaturation Curves. In Shirley, B. A. (Ed) Protein stability and folding. Theory and Practice. Methods in Molecular Biology. **40**, Human Press Inc. pp 177-190

Singh, S.M., Panda, A.K.,(2005). Solubilisation and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. **99**: 303–310.

Slinde,E., and Flatmark, T. (1976). Effect of the hydrophile-lipophile balance of non-ionic detergents (Triton X-series) on the solubilization of biological membranes and their integral *b*-type cytochromes. Biochim. Biophys. Acta **455**: 796-805.

Speed, M.A., Wang, D.I.C., King, J., (1996). Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. Nat. Biotechnol. **14**: 1283–1287.

Srinivas, V. R., Bhanuprakash Reddy, G., Ahmad, N., Swaminathan, C. P., Mitra N., and Surolia A. (2001) Legume lectin family, the 'natural mutants of quaternary state', provide insights into the relationship between protein stability and oligomerisation. Biochim. Biophys. Acta. **1527**: 102-111).

Stempfer, G., HollNeugebauer, B., Kopetzki, E., Rudolph, R., (1996). A fusion protein designed for noncovalent immobilization: stability, enzymatic activity, and use in an enzyme reactor. Nat. Biotechnol. **14**: 481–484.

Studier, F.W. and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**, 113–130.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., & DuBendorff, J.W. (1991). Use of T7 RNA polymerase to direct expression of cloned genes. Methods of Enzymology. **188**: 60

Sumner, J. B. (1919) The globulins of the jack bean Canavalia ensiformis J. Biol. Chem. **37:** 137-142

Sumner, J. B. and Howell, S. F. (1936). 'The identification of the hemagglutinin of the jack bean with Concanavalin A'. J. Bacteriol. **32:** 227-37

Terashima, M., Suzuki, K., and Katoh, S. (1996). Effective refolding of fully reduced lysozyme with a flow-type reactor. Process Biochem. **31**: P. 341-345:

Troganis, A. (1992). PhD thesis, Univ. of Athens, Greece.

Tsumoto, K. , Shinoki, K., Kondo, H., Uchikawa, M., Juji, T., Kumagai, I., (1998). Highly efficient recovery of functional single-chain Fv fragments from inclusion bodies overexpressed in Escherichia coli by controlled introduction of oxidizing reagents—application to a human single-chain Fv fragment, J. Immunol. Methods. **219**:119–129.

Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Arakawa, T.,(2003). Solubilisation of active green fluorescent protein from insoluble particles by guanidine and arginine. Biochem. Biophys. Res. Commun. **312**: 1383–1386.

Vallejo LF, Rinas U. (2004). Optimised procedure for renaturation of recombinant human bone morphogenetic protein-2 at high protein concentration. Biotechnol Bioeng.**85**:601–609.

Ventura, S., and Villaverde, A., (2006). Protein quality in bacterial inclusion bodies. TRENDS in Biotechnology, **24**:179-183.

Vincent, P., Dieryck, W., Maneta-Peyret, L., Moreau, P., Cassagne, C., Santarelli, X., (2004). Chromatographic purification of an insoluble histidine tag recombinant Ykt6p SNARE from Arabidopsis thaliana overexpressed in E. coli, J. Chromatogr. B **808**: 83–89.

West SM, Chaudhuri JB, Howell JA. (1998). Improved protein refolding using hollow-fibre membrane dialysis. Biotechnol Bioeng.**57**:590–599.

Wetlaufer DB, Xie Y. (1995). Control of aggregation in protein refolding: a variety of surfactants promote renaturation of carbonic anhydrase II. Protein Sci, **4**:1535-1543.

Xiao-Yan Dong, Jian-Hua Liu, Fu-Feng Liu, Yan Sun, (2007). Self-interaction of native and denatured lysozyme in the presence of osmolytes, l-arginine and guanidine hydrochloride Biochemical Engineering Journal, **In Press, Accepted Manuscript**, Available online 5 November 2008

Xie Y, Wetlaufer D.B., (1996). Control of aggregation in protein folding: the temperature-leap tactic. Protein Sci, **5**:517-523.

Yansura, D. G.,and Henner, D.J., (1990). Use of *Escherichia coli trp* promoter for direct expression of proteins. Methods Enzymol, **185**: 54-60.

Yoshii H, Furuta T, Yonehara T, Ito D, Linko Y-Y, Linko P. (2000). Refolding of denatured/reduced lysozyme at high concentration with diafiltration. Biosci Biotechnol Biochem.; **64:1159**–1165.

Zardeneta G, Horowitz P.M., (1994). Detergent, liposome, and micelle assisted protein refolding. Anal Biochem, **223:**1-6.

Zardeneta G, Horowitz P.M., (1994). Protein refolding at high concentrations using detergent/phospholipid mixtures. Anal Biochem, **218:**392-398.